第一推动丛书·综合系列

（套装共8册）



- 不可错过的综合科学经典
- 《第一推动丛书》25周年全新集结
- 涉及未来图景、人工智能、复杂系统等内容
- 多位诺贝尔科学家呕心沥血作稿
- 世界公认有创见的思想家成书

湖南科学技术出版社

VISIT…

LANZAROTE Caliente.com

目　录

复杂
复杂的引擎
皇帝新脑
逻辑的引擎
数学的意义
未来50年
真理与美
下一步是什么

第一推动丛书: 综合系列

The Polytechnique Series

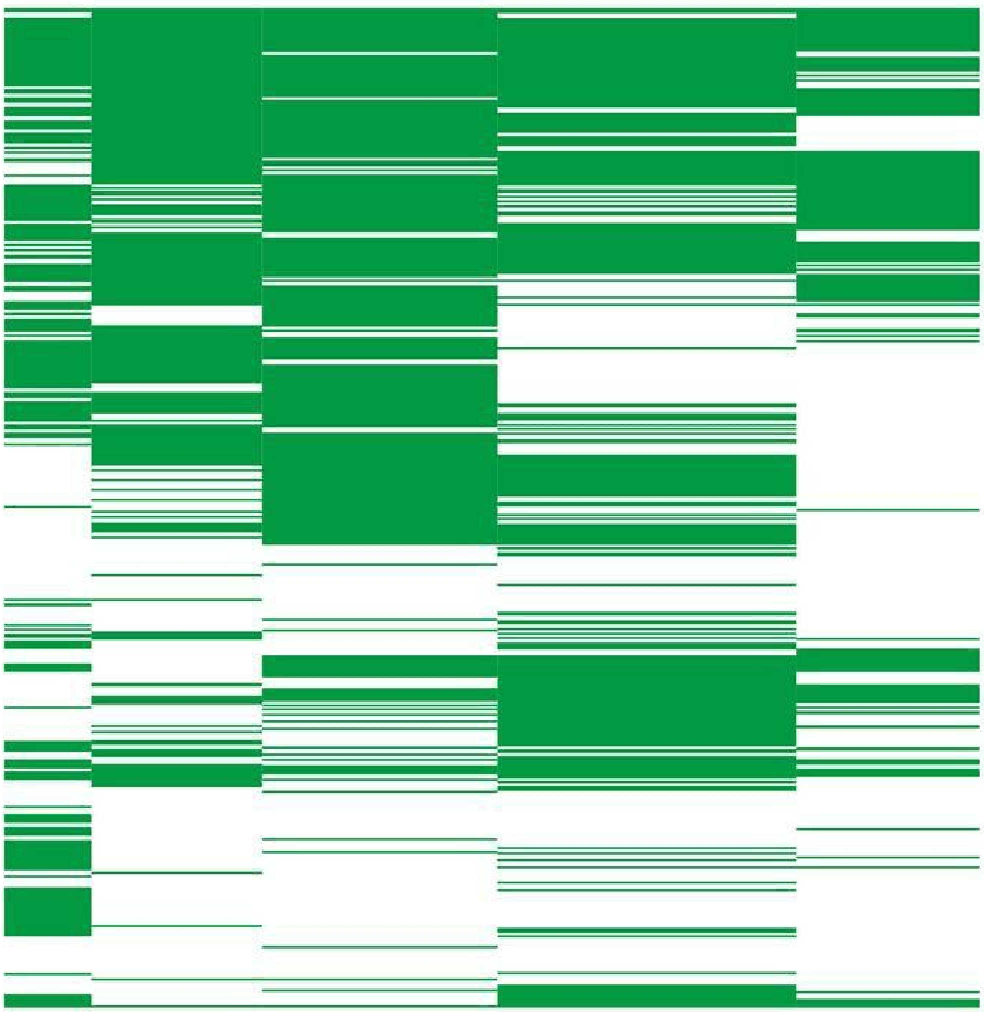
复杂

Complexity

THE FIRST MOVER

[美] 梅拉妮·米歇尔 著
唐璐 译

Melanie Mitchell



湖南科学技术出版社

版权信息

复杂

著者：（美）梅拉妮·米歇尔

译者：唐璐

责任编辑：吴炜　戴涛　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794369

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
前言
致谢
1 背景和历史
第1章 复杂性是什么
昆虫群落
大脑
免疫系统
经济
万维网
复杂系统的共性
如何度量复杂性
第2章 动力学、混沌和预测
动力系统理论的起源
对预测的重新认识
线性兔子和非线性兔子
逻辑斯蒂映射
混沌的共性
混沌思想带来的革命
第3章 信息
信息是什么
能量、功、熵
麦克斯韦妖
统计力学提要
微观态与宏观态
香农信息
第4章 计算
什么是计算？什么可以计算
希尔伯特问题和哥德尔定理
图灵机和不可计算性
图灵机概述

一个简单的例子
定义为图灵机的明确程序
通用图灵机
图灵对判定问题的解决
哥德尔和图灵的命运
第5章 进化
达尔文之前的进化观念
达尔文理论的起源
孟德尔和遗传律
现代综合
对现代综合的挑战
第6章 遗传学概要
第7章 度量复杂性
用大小度量复杂性
用熵度量复杂性
用算法信息量度量复杂性
用逻辑深度度量复杂性
用热力学深度度量复杂性
用计算能力度量复杂性
统计复杂性
用分形维度量复杂性
用层次性度量复杂性
2 计算机中的生命和进化
第8章 自我复制的计算机程序
生命是什么
计算机中的自我复制
自我复制程序的深层意义
DNA的自我复制
冯·诺依曼的自复制自动机
冯·诺依曼
第9章 遗传算法
遗传算法菜谱
遗传算法的应用
进化的罗比，易拉罐清扫机器人
GA演化的策略是如何解决这个问题的
GA是如何演化出好的技巧的

3 大写的计算
第10章 元胞自动机、生命和宇宙
自然界中的计算
元胞自动机
生命游戏
四类元胞机
沃尔夫勒姆的“新科学”
第11章 粒子计算
第12章 生命系统中的信息处理
什么是信息处理
传统计算机中的信息处理
元胞自动机中的信息处理
免疫系统
蚁群
生物代谢
这些系统中的信息处理
第13章 如何进行类比（如果你是计算机）
容易的事很难
进行类比
我对类比的认识经历
简化的类比
模仿者
如何做到
模仿者程序
运行模仿者
总结
第14章 计算机模型
模型是什么
理想模型
对合作的进化进行模拟
囚徒困境
建模的好处
计算机建模注意事项
4 网络
第15章 网络科学
小世界

网络新科学
什么是网络思维
到底什么是“网络”
小世界网络
无尺度网络
万维网的度分布
无尺度分布和钟形曲线
网络稳健性
第16章 真实世界中的网络
真实世界中的网络
基因调控网络
网络思想的意义
无尺度网络是如何产生的
幂律以及对其的质疑
网络中的信息传播和连锁失效
第17章 比例之谜
生物学中的比例缩放
一次跨学科合作
幂律与分形
代谢比例理论
理论的应用
争议
幂律的未解之谜
第18章 进化，复杂化
遗传，复杂化
基因是什么
进化发育生物学
基因调控和考夫曼的“秩序的起源”
对考夫曼的研究的反响
总结
5 尾声
第19章 复杂性科学的过去和未来
统一理论和一般性原理
复杂系统研究的根源
五个问题
复杂性的未来，等待卡诺

附录 访谈——梅拉妮·米歇尔谈复杂性
参考文献

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

献给侯世达和霍兰德

前言

还原论[1]是对这个世界最自然的理解方式。它是说“如果你理解了整体的各个部分，以及把这些部分‘整合’起来的机制，你就能够理解这个整体”。只要是精神正常的人就不会反对还原论。

——侯世达（Douglas Hofstadter），

《哥德尔、艾舍尔、巴赫——集异璧之大成》

从17世纪以来，还原论就一直在科学中占据着主导地位。还原论最早的倡议者之一笛卡儿这样描述他的科学方法：“将面临的所有问题尽可能地细分，[2]细至能用最佳的方式将其解决为止”，并且“以特定的顺序引导我的思维，从最简单和最容易理解的对象开始，一步一步逐渐上升，直至最复杂的知识”。

从笛卡儿、牛顿等现代科学奠基者的时代，直到20世纪初，科学的主要目标都是用基础物理学来对一切现象进行还原论式的解释。19世纪末许多科学家都赞同物理学家迈克耳孙1894年说的一句名言：“大部分大的基本原理似乎[3]已经被明确建立起来了，今后的进展主要是将这些原理严格应用到值得我们注意的一些现象中去。”

此后的30年里，物理学又有了相对论和量子力学这样革命性的发现。但20世纪的科学也见证了还原论梦想的破灭。虽然基础物理学和还原论对于解释极大和极小的事物取得了伟大的成就，但在对于接近人类尺度的复杂现象的解释上，它们却保持惊人的沉默。

还原论的计划在许多现象面前都止步不前：天气和气候似乎无法还原的不可预测性；生物以及威胁它们的疾病的复杂性和适应性；社会的经济、政治和文化行为；现代技术与通信网络的发展和影响；智能的本质以及用计算机实现智能的可能前景。对复杂行为如何从简单个体的大规模组合中出现进行解释时，混沌、系统生物学、进化经济学和网络理论等新学科胜过了还原论，反还原论者的口号——“整体大于部分之

和”——也随之变得越来越有影响力。

20世纪中叶，许多科学家意识到，这类现象无法被归入单个学科，而需要在新的科学基础之上从交叉学科的角度进行理解。一些人开始尝试建立新的基础，这其中包括控制论、协同学、系统科学，以及最近才出现的——复杂系统科学。

1984年，来自不同学科的24位科学家和数学家在新墨西哥州圣塔菲的高原沙漠会聚一堂，讨论“科学中涌现的综合”。[4] 他们的目标是筹划建立一家新的研究机构，“致力于研究各种高度复杂和相互作用的系统，这些系统只有在交叉学科的背景下才能研究清楚”并“推动知识的统一和共担责任的意识，[5] 与目前盛行的知识界的各自为政作斗争”。就这样，圣塔菲研究所作为复杂系统的研究中心被建立起来了。

1984年我还没有听说过“复杂系统”一词，虽然头脑中已经有了类似的想法。我当时是密歇根大学计算机系的一年级研究生，研究方向是人工智能，也就是让计算机像人一样思维。事实上，我的一个目标就是理解人类如何思维——万亿个微小的脑细胞以及它们的电和化学通信如何涌现出抽象思维、情感、创造性，甚至意识。我曾深深迷恋于物理学和还原论的目标，后来才领悟到，目前的物理学对于智能可以做的很少，即便是专门研究大脑细胞的神经科学，也无法理解思维如何从大脑活动中涌现出来。很显然还原论者对认知的研究是误入歧途——我们根本无法在单个神经元和突触的层面上理解认知。

因此，虽然我以前没有听说过“复杂系统研究”，它却很快引起了我的强烈共鸣。同时我也感到，我自己的研究领域——计算机科学——在这里可以大有作为。受研究计算的先驱们影响，我觉得计算的思想要比操作系统、编程语言、数据库之类的东西深刻得多，计算的本质与生命和智能的内在本质有密切的关联。我很幸运，在密歇根大学，“自然系统中的计算”是系里的核心课程，与软件工程和编译器设计一样。

1989年，我攻读研究生的最后一年，我的博士生导师侯世达受邀参加在新墨西哥州洛斯阿拉莫斯举行的主题为“涌现计算”的研讨会。[6] 他太忙了抽不出时间，因此就让我替他去。在这样高水平的会议上报告自己的工作让我既兴奋又害怕。就是在这次会议上，我第一次遇见了一大群和我抱有同样想法的人。我发现他们不仅为这样的想法取了一个名字

——复杂系统——而且他们在圣塔菲附近的研究所正是我想去的地方。我决定在这里争取一个职位。

不断坚持，再加上运气，我终于获得了圣塔菲研究所（Santa Fe Institute）的邀请，在那里访问一个夏天。一个夏天又延长为一年，后来又延长了一年。最终我成为研究所的常驻研究人员。来自不同国家和学科的人们聚集在这里，一起从不同的角度来探索同样的问题。我们如何超越还原论的传统范式，对似乎无法还原的复杂系统形成新的理解？

这本书源自我为圣塔菲的乌拉姆纪念讲座（Ulam Memorial Lecture）做的演讲——这个讲座为普通听众举办，是关于复杂系统的年度系列讲座，以纪念伟大的数学家乌拉姆。我的系列演讲的题目是“复杂性科学的过去和未来”。要为非专业听众讲清楚领域广泛的复杂性研究，让他们理解研究的现状和广阔的前景，这极具挑战性。我的角色很像是在一个幅员辽阔、文化多样的异国的导游。我们只有很短的时间来了解历史背景，参观著名景点，并感受这里的风土人情，必要时还要进行翻译以便于理解。

这本书就是由这些讲座扩充而成——就像观光指南。书中讲述的是让我也让研究复杂系统的其他人曾经或正在着迷的问题：自然界中我们认为复杂和具有适应性的系统——大脑、昆虫群落、免疫系统、细胞、全球经济、生物进化——如何通过简单规则产生出复杂和适应性的行为？相互依赖而又自私的生物是如何一起协作，以解决影响它们整体生存的问题？这些现象存在普遍规律吗？生命、智能和适应性能用机械和计算实现吗？如果能，我们又能不能建造出真正具有生命和智能的机器？如果能做到，我们又应不应该这样做呢？

我听说随着学科间的界线变得模糊，科学术语的意义也会变得模糊。研究复杂系统的人们谈论各种模糊的概念，例如自发秩序、自组织、涌现（包括“复杂性”本身）。这本书的一个主要目的就是为这些人所谈论的提供一幅清晰的图景，并探讨这些交叉学科的概念和方法是否能产生出实用的科学和新的思想，以解决人类面临的各种难题，例如疾病的传播、世界自然和经济资源分配的不公平、武器扩散和冲突的增多，以及人类社会对环境和气候的影响。

这本书就像一本复杂性科学的核心思想的观光指南——它们从何而

来，又将到哪里去——再加上我自己的一些见解。对于正在发展中的科学领域，其核心思想、意义以及可能导致的后果，人们的认识会（略）有不同。因此我的观点与其他专家也许会不一样。本书中一个重要的部分就是阐释这些差别，另外我也将尽我所能介绍一下那些未知的或刚刚开始被理解的领域。正是这些使得科学引人入胜，值得去探索和了解。我希望能让读者也感受到这些思想的迷人魅力和探索它们的过程中那种无可比拟的兴奋感觉。

本书分为5部分。在第1部分我将介绍4个主题的历史和内容，这4个主题是复杂系统研究的基础：信息、计算、动力学和混沌、进化。在第2到第4部分我将阐述这4个主题如何在复杂性科学中被组织到一起。我将描述如何在计算机中模拟生命和进化，以及计算的概念反过来又如何被用来解释自然系统的行为。我还会介绍网络科学的发展，以及网络科学发现的社会群体、互联网、传染病和生物代谢等各种系统中存在的深刻共性。另外，我还会用各种例子说明如何测量自然界中的复杂性，它又如何改变我们对生命系统的认识，以及这些新的认识能不能引导智能机器的设计。我会介绍复杂系统的各种计算机模型，以及这些模型所面临的风险。最后，书的末尾还将讨论寻找复杂性科学一般性原则的问题。

要理解书中内容无需数学或科学的背景知识，在涉及的时候我会小心地循序渐进。我希望这本书对专家和非专业读者都会有价值。虽然讨论不是技术性的，但我还是会尽力做到言而有物。注释中给出了引文的出处和讨论的附加内容，以及为想深入学习的读者准备的科学文献索引。

你对复杂性科学感到好奇吗？想不想探索一番呢？让我们出发吧。

致谢

受圣塔菲研究所（SFI）邀请主持复杂系统暑期学校和为乌拉姆纪念讲座演讲的经历激发了我写这本书的念头，在此向圣塔菲研究所表示感谢。同时也要感谢SFI多年来为我提供了极具启发而且富有成效的科学氛围。SFI大家庭中的众多科学家慷慨地分享了他们的思想，给了我很多灵感，这里无法将他们一一列举，在此向他们全体表示感谢。还要感谢SFI的工作人员，我在研究所工作期间，他们真诚友善地给予了我帮助。

感谢Bob Axelrod、Liz Bradley、Jim Brown、Jim Crutchfield、Doyne Farmer、Stephanie Forrest、Bob French、Douglas Hofstadter、John Holland、Greg Huber、Ralf Juengling、Garrett Kenyon、Tom Kepler、David Krakauer、Will Landecker、Manuel Marques-Pita、Dan Mc Shea、John Miller、Jack Mitchell、Norma Mitchell、Cris Moore、David Moser、Mark Newman、Norman Packard、Lee Segel、Cosma Shalizi、Eric Smith、Kendall Springer、J.Clint Sprott、Mick Thomure、Andreas Wagner、Chris Wood。他们热情地为我答疑，对书稿提出意见，并帮助我对书的内容有了更清晰的认识。当然，如果书中有任何不当之处，作者文责自负。

还要感谢牛津的编辑Kirk Jensen和Peter Prescott对我自始至终的支持和超凡的耐心，以及牛津的Keith Faivre和Tisse Takagi给予的帮助。感谢谷歌学术、谷歌图书、亚马逊网站以及经常不怎么公道但又极为有用的维基百科，它们使得学术搜索变得极为便利。

我要将这本书献给侯世达和霍兰德，他们对我的工作和生活给予了如此多的启发和鼓励，能得到他们的教诲和友爱是我三生有幸。

最后，感谢我的家人：我的父母Jack和Norma Mitchell、兄弟Jonathan Mitchell以及我的丈夫Kendall Springer，感谢他们给予我的爱和支持。感谢Jacob和Nicholas Springer，虽然他们的到来延误了这本书的写作，但他们也给我们的生活带来了新的欢乐和惹人喜爱的复杂性。

1　背景和历史

科学已经探索了微观和宏观世界；[7] 我们对所处的方位已经有了很好的认识。亟待探索的前沿领域就是复杂性。

——斐杰斯（Heinz Pagels），

《理性之梦》（Te Dreams of Reason ）

第1章　复杂性是什么

一些思想是由简单的思想组合而成，[8] 我称此为复杂；比如美、感激、人、军队、宇宙等。

——洛克（John Locke），

《人类理解论》（An Essay Concerning Human Understanding ）

巴西：亚马孙雨林。几十万只行军蚁[9]（army ant）在行进。没有谁掌控这支军队，不存在指挥官。单个蚂蚁几乎没有什么视力，也没有多少智能，但是这些行进中的蚂蚁聚集在一起组成了扇形的蚁团，一路风卷残云，吃掉遇到的一切猎物。不能马上吃掉的就会被蚁群带走。在行进了一天并摧毁了足球场大小的浓密雨林中的一切食物后，蚂蚁会修筑夜间庇护所——由工蚁连在一起组成的球体，将幼蚁和蚁后围在中间保护起来。天亮后，蚁球又会散成一只只蚂蚁，各就各位进行白天的行军。

专门研究蚂蚁习性的生物学家弗兰克斯（Nigel Franks）写道，“单只行军蚁[10]是已知的行为最简单的生物”，“如果将100只行军蚁放在一个平面上，它们会不断往外绕圈直到体力耗尽死去”。然而，如果将上百万只放到一起，群体就会组成一个整体，形成具有所谓“集体智能（collective intelligence）”的“超生物（superorganism）”[11]。

这究竟是怎么回事呢？虽然科学家们已经很熟悉蚁群的习性，但集体智能的产生机制依然是个谜。就像弗兰克斯所说，“我研究了布氏游蚁[12]（E.burchelli，一种常见的行军蚁）很多年，我发现，对它们的社会结构了解得越多，对其社会组织的疑问就会越多”。

行军蚁是许多我们认为“复杂”的自然和社会系统的缩影。蚂蚁、白蚁以及人类这样的社会生物会聚集在一起，共同形成复杂的社会结构，从而增加种群整体的生存机会，目前还没有人确切地知道其背后的机理。类似的还有，免疫系统如何抵抗疾病，细胞如何自组织成眼睛和大脑，经济系统中自利的个体如何形成结构复杂的全球市场。最为神秘的是，所谓的“智能”和“意识”是如何从不具有智能和意识的物质中涌现出来的。

这些正是复杂系统所关注的问题。复杂系统试图解释，在不存在中央控制的情况下，大量简单个体如何自行组织成能够产生模式、处理信息甚至能够进化和学习的整体。这是一个交叉学科研究领域。复杂一词源自拉丁词根plectere，意为编织、缠绕。在复杂系统中，大量简单成分相互缠绕纠结，而复杂性研究本身也是由许多研究领域交织而成。复杂系统专家认为，自然界中的各种复杂系统——比如昆虫群落、免疫系统、大脑和经济——之间，具有许多共性。下面我们来一一了解。

昆虫群落

社会性昆虫群落提供了极为丰富而神奇的复杂系统范例。例如，一个蚁群可能由数百只乃至上百万只蚂蚁组成，单只蚂蚁其实都相对简单，它们受遗传天性驱使寻找食物，对蚁群中其他蚂蚁释放的化学信号做出简单反应，抵抗入侵者，等等。然而，任何一个在野外观察过蚁群的人都会意识到，虽然单只蚂蚁的行为很简单，但整个蚁群一起构造出的结构却复杂得惊人，而且这种结构明显对群体的生存极为重要。它们使用泥土、树叶和小树枝建造出极为稳固的巢穴，巢穴中有宏大的通道网络，育婴室温暖而干爽，温度由腐烂的巢穴材料和蚂蚁自身的身体控制。一些种类的蚂蚁还会将它们的身体相互连在一起组成很长的桥，从而可以跨越很长的距离（对它们来说很长），通过树干转移到另一蚁穴（图1.1）。科学家们对蚂蚁及其社会结构进行了细致的研究，但现在仍然无法彻底弄清它们的个体和群体行为：蚂蚁的个体行为如何形成庞大而复杂的结构，蚂蚁之间如何相互通信，蚁群作为整体如何适应环境

变化（比如天气变化和受到攻击）。生物进化又是如何产生出个体如此简单、整体上却如此复杂的生物？



图1.1　蚂蚁用身体建造出一座桥，让蚁群能迅速通过沟壑（图片由Carl Rettenmeyer提供）

大脑

认知科学家侯世达在《哥德尔、艾舍尔、巴赫——集异璧之大成》一书中 [13] 对蚁群和大脑进行了比较。两者都是由相对简单的个体组成，个体之间只进行有限的通信，整体上却表现出极为复杂的系统（“全局”）行为。在大脑中，简单个体是神经元。除了神经元，大脑中还有许多不同的细胞，但绝大多数脑科学家都认为是神经元的活动以及神经元群的连接模式决定了感知、思维、情感、意识等重要的宏观大脑活动。

图1.2（上图）就是神经元的图像。神经元主要由三部分组成：细胞体，接收其他神经元信号的分支（树突），以及向其他神经元发送信号的主干（轴突）。大致上，神经元可以处于活跃状态（激发）或非活跃状态（未激发）。当神经元通过树突从其他神经元接收到足够强的信号时，它就会激发。激发时会通过轴突传出电信号，然后释放出神经递质转换成化学信号，化学信号又会作用于其他神经元的树突对其进行触发。神经元的激发频率和产生的化学输出信号会根据输入和最近的激发状况随时间变化。

这与蚁群很类似：个体（神经元或蚂蚁）之间相互传递信号，信号的总强度达到一定程度时，会导致个体以特定的方式动作，从而再次产生信号。总体上会产生非常复杂的效果。前面说过对蚂蚁及其社会结构尚未完全了解；同样，对于单个神经元的行为和庞大的神经网络如何产生出大脑的宏观行为（图1.2，下图），科学家们也没有弄清楚。他们不知道神经元信号的意义，不知道大量神经元如何一起协作产生出整体上的认知行为，也不知道它们是怎样让大脑能够思维和学习新事物。同样，最让人迷惑的也许就是，如此精巧、整体能力如此强大的信号系统是怎样进化出来的。

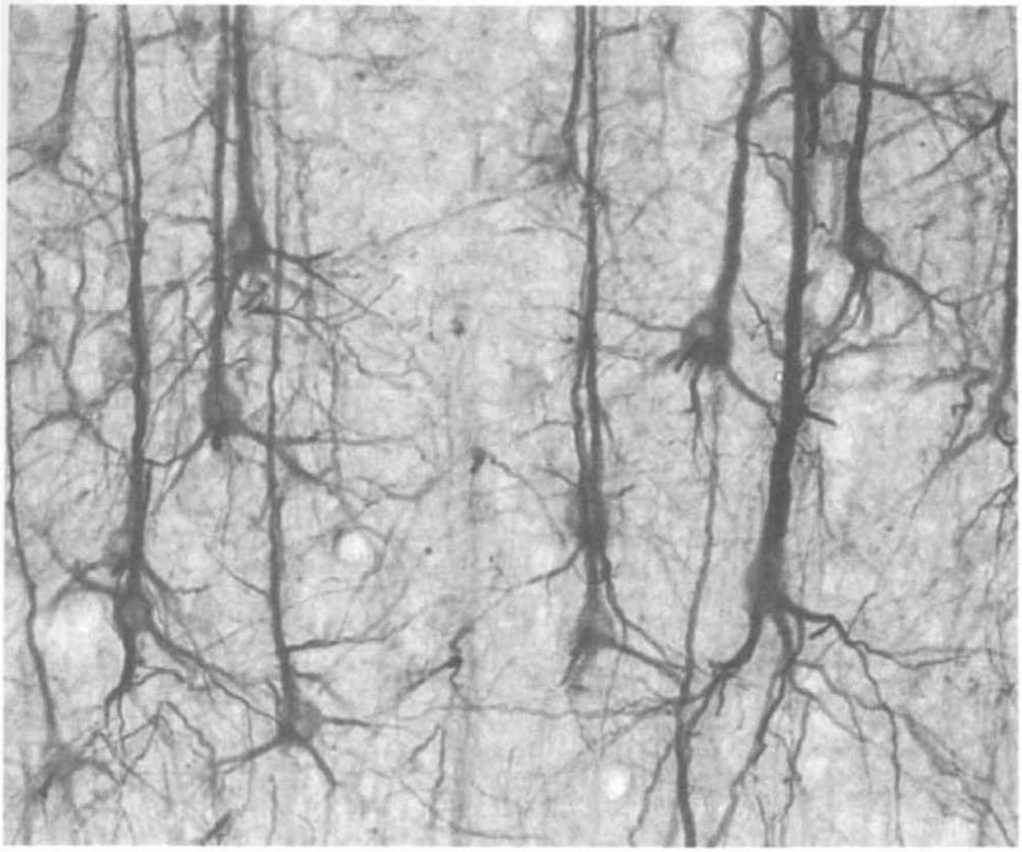




图1.2　上图：神经元着色显微图像。下图：人类大脑。一个层次上的行为是如何产生出更高层次上的行为呢？[神经元图像来自brainmaps.org（http：//brainmaps.org/smi32-pic.jpg），由知识共享组织（Creative Commons）授权使用（http：//creativecommons.org/licenses/by/3.0/）。大脑图像由Christian R.Linder提供]

免疫系统

免疫系统是又一个例子。在免疫系统中，相对简单的组分一起产生出包含信号传递和控制的复杂行为，并不断进行适应。图1.3展现了免疫系统的复杂性。

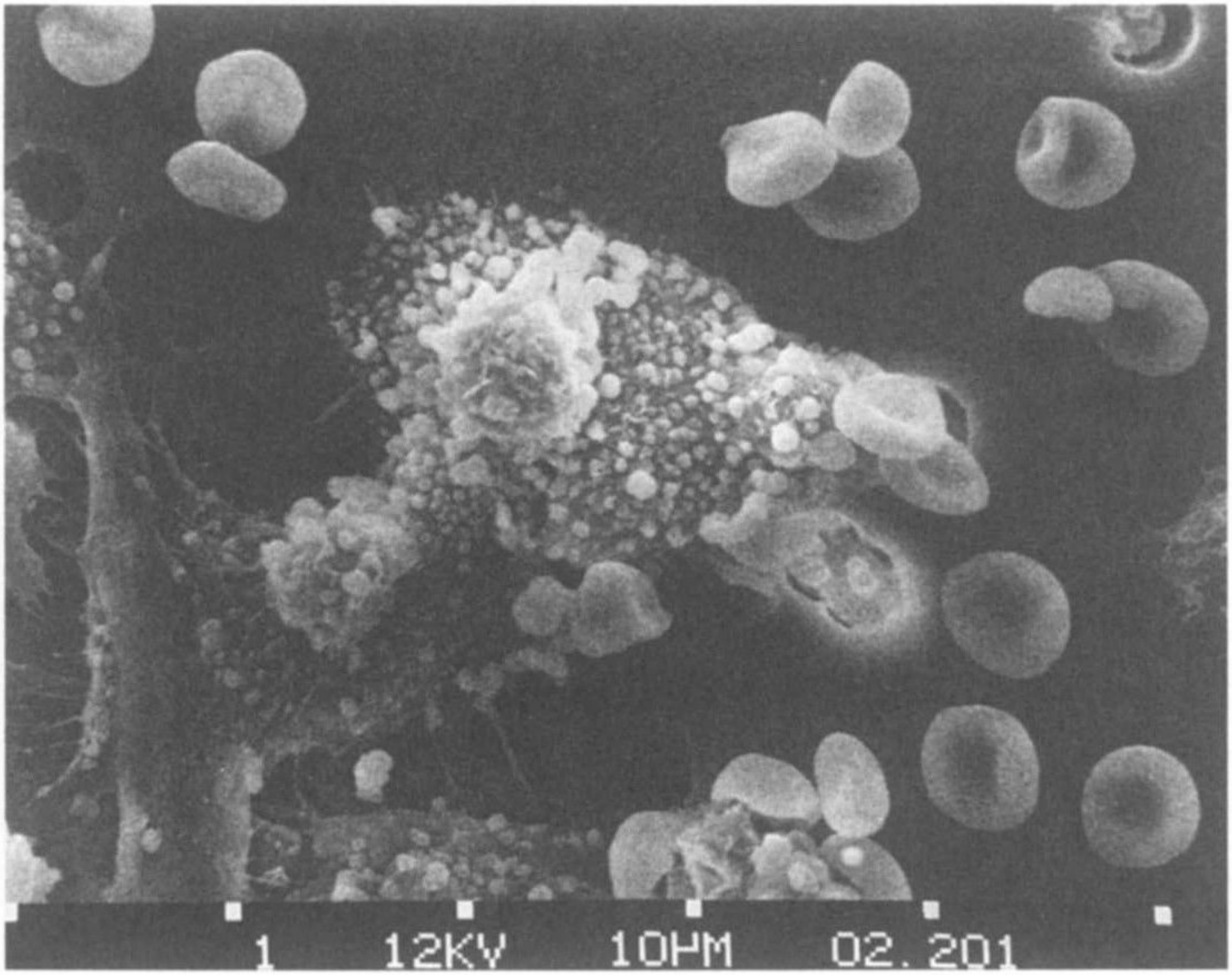


图1.3　免疫细胞攻击癌症细胞[Susan Arnold摄影，图片来自美国国家癌症研究所视觉在线网站（National Cancer Institute Visuals Online）（http：//visualsonline.cancer.gov/details.cfm?imageid=2370）]

同大脑一样，不同动物的免疫系统的复杂程度也各不相同，但总体上的原则是一样的。免疫系统由许多不同的细胞组成，分布在身体各处（血液、骨髓、淋巴结等）。这些细胞在没有中央控制的情况下一起高效地工作。

免疫系统中的主角是白细胞，也称为淋巴细胞。白细胞能通过其细胞体上的受体识别与某种可能入侵者（比如细菌）相对应的分子。大量白细胞哨兵在血液中不停巡逻，如果被激活——也就是特定受体偶然遇到了与其匹配的入侵者——就发出警报。一旦淋巴细胞被激活，就会分泌出大量能够识别类似入侵者的分子——抗体。这些抗体会到处去搜寻和摧毁入侵者。被激活的淋巴细胞的分裂速度也会加快，从而产生出更多后代淋巴细胞，帮助搜寻入侵者和释放抗体。后代淋巴细胞会不断繁

衍，从而让身体能记住入侵者特征，再次遇到这种入侵者时就能具有免疫力。

有一类细胞被称为B细胞（B是指它们产生自骨髓，Bone marrow），它具有一种奇特的性质：B细胞与某种入侵者匹配得越好，它产生的后代细胞就越多。通过随机变异，子细胞与母细胞会稍有不同，而这些子细胞产生后代的能力也与它们同入侵者相匹配的程度成正比。这样就形成了达尔文自然选择机制，B细胞变得与入侵者越来越匹配，从而产生出能极为高效地搜寻和摧毁微生物罪犯的抗体。

还有许多种类的细胞也参与了免疫反应的大合奏。T细胞（产生自胸腺，Thymus）对于调节B细胞的反应很重要。巨噬细胞四处游荡，寻找已被抗体标记的东西，然后将其摧毁。有些细胞让免疫能长期有效。此外，系统中还有一部分是用来防止免疫系统攻击身体的正常细胞。

同大脑和蚁群一样，免疫系统的行为是通过大量简单参与者的独自行动产生，并没有谁在进行掌控。简单参与者——B细胞、T细胞、巨噬细胞，等等——的行动可以看作某种化学信号处理网络，一旦有一个细胞识别出入侵者就会触发细胞之间产生信号雪崩，从而产生精巧而复杂的反应。不过目前这个信号处理系统的许多关键细节还没有研究清楚。比如，目前仍然没有完全弄清楚相关的信号是什么，它们具体的功能是什么，它们又是如何相互协作，从而使得系统作为一个整体能够“知道”环境中存在何种威胁，并产生出应对这种威胁的长期免疫力。我们也不清楚这种系统是如何避免攻击身体；又是什么导致系统失灵，例如如果患有自身免疫病（autoimmune diseases），系统就会对身体发起攻击；艾滋病毒（HIV）又是用怎样的策略直接攻击免疫系统本身。同样，还有一个关键问题，就是这样高效的复杂系统当初是如何进化出来的。

经济

经济也是复杂系统，在其中由人（或公司）组成的“简单、微观的”个体购买和出售商品，而整个市场的行为则复杂而且无法预测，比如不同地区的住宅价格或股价的波动（图1.4）。很多经济学家认为经济在微观和宏观层面上都具有适应性。在微观层面上，个人、公司和市场都试图通过研究其他人和公司的行为来增加自己的收益。以前一直认

为，微观上的自利行为会使得市场在总体上——宏观层面上——趋于均衡，在均衡状态下商品价格无论怎样变化都无法让所有人受益。从收益或消费者满意度来看，如果有人受益，就肯定会有人受损。市场能达到均衡态就认为市场是有效的。18世纪经济学家亚当·斯密（Adam Smith）将市场的这种自组织行为称为“看不见的手”：它产生自无数买卖双方的微观行为。




道琼斯工业平均指数，1970~2008
收盘价
14000
12000
10000
8000
6000
4000
2000
0
1972年1月
1982年1月
1992年1月
2002年1月
月份

图1.4　个体的交易行为产生出金融市场无法预测的宏观行为。上图：纽约股票交易所[图片来自纽约公立图书馆麦斯坦部（Milstein Division of US History, Local History and Genealogy, The New York Public Library），经阿斯特、莱诺克斯和狄尔登基金（Astor, Lenox and Tilden Foundations）许可使用]。下图：1970～2008年各月道琼斯工业平均指数收盘价

经济学家感兴趣的问题是，市场怎样才会变得有效，以及反过来，为何在现实世界中市场会失效。近年来，关注复杂系统研究的经济学家开始尝试用复杂系统的术语来解释市场的行为：动力学无法预测的全局行为模式，比如市场泡沫及其崩溃的模式；信号和信息的处理，比如个体买卖者的决策过程，以及市场作为整体“计算”有效价格的“信息处理”能力；还有学习和适应，比如商家调整产品以适应消费者的需求变化，以及市场作为一个整体对价格进行调整。

万维网

万维网诞生于20世纪90年代初，此后呈爆炸性增长。与前面描述的系统类似，万维网可以视为自组织的社会系统：每个人都看不到网络的全貌，只是简单地发布网页并将其链接到其他网页。然而，复杂系统专家发现这个网络在整体上具有一些出人意料的宏观特性，包括其结构、增长方式，信息如何通过链接传播，以及搜索引擎和万维网链接结构的协同演化，这一切都可以视为系统作为一个整体的“适应”行为。万维网从简单规则中涌现出的复杂行为是目前复杂系统研究的热点。图1.5展现了一部分网页以及其链接的结构。似乎许多部分都很相似，问题是，为什么会这样？

复杂系统的共性

这些系统在细节上很不一样，但如果从抽象层面上来看，则会发现它们有很多有趣的共性。

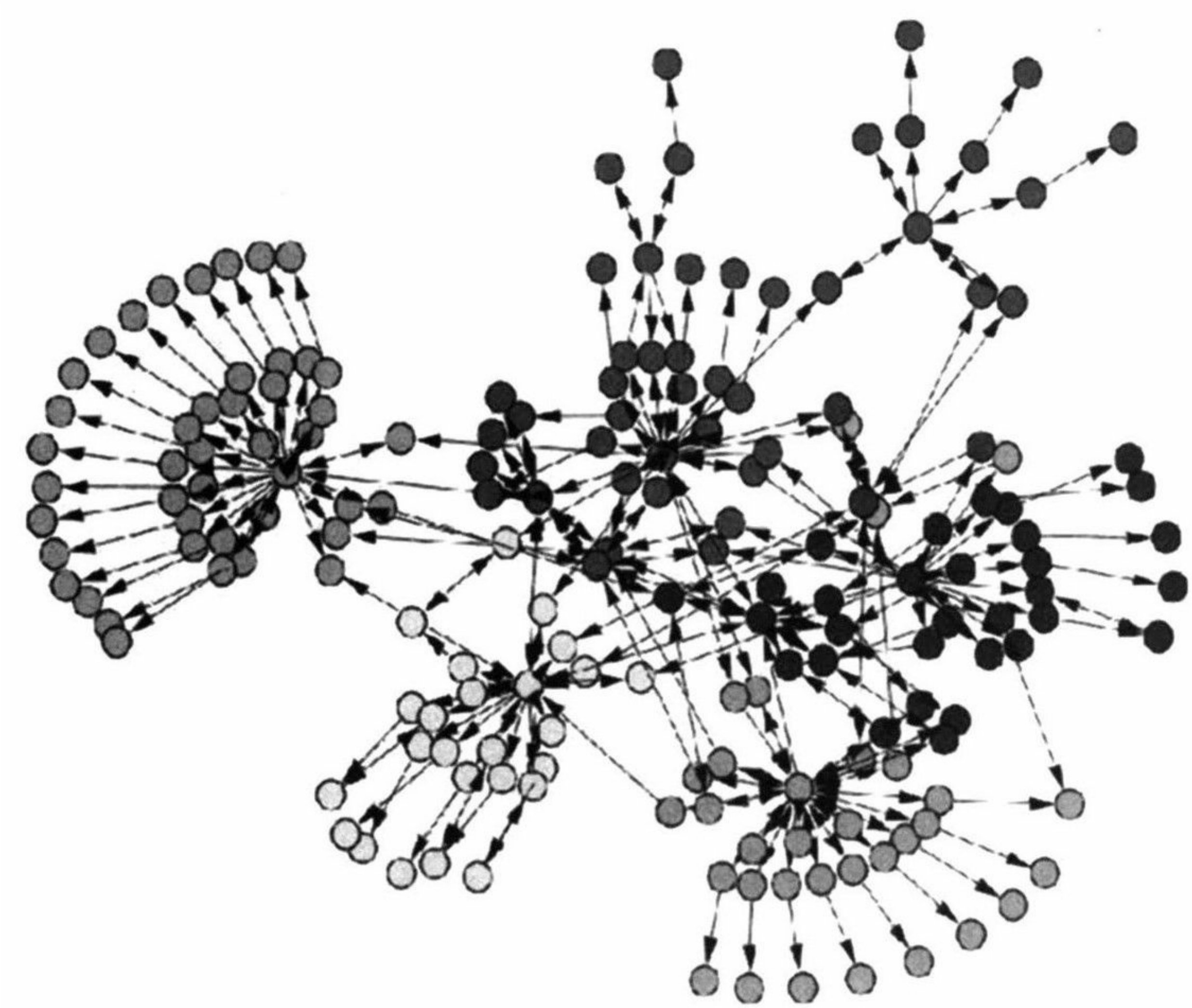


图1.5　万维网的部分网络结构（引自M.E.J.Newman&M.Girvin, Physical Review Letters E，69，026113，2004。美国物理学会版权所有。经许可重印）

1.复杂的集体行为：前面讲到的所有系统都是由个体组分（蚂蚁、B细胞、神经元、股票交易者、网站设计人员）组成的大规模网络，个体一般都遵循相对简单的规则，不存在中央控制或领导者。大量个体的集体行为产生出了复杂、不断变化而且难以预测的行为模式，让我们为之着迷。

2.信号和信息处理：所有这些系统都利用来自内部和外部环境中的信息和信号，同时也产生信息和信号。

3.适应性：所有这些系统都通过学习和进化过程进行适应，即改变自身的行为以增加生存或成功的机会。

现在我可以对复杂系统加以定义：复杂系统是由大量组分组成的网络，不存在中央控制，通过简单运作规则产生出复杂的集体行为和复杂的信息处理，并通过学习和进化产生适应性。[有时候会对复杂适应系统（在其中适应性扮演重要角色）和复杂非适应系统（比如飓风或湍流）加以区分。在书中讨论的大部分系统都是适应性的，我不再区分。]

如果系统有组织的行为不存在内部和外部的控制者或领导者，则称之为自组织（self-organizing）。由于简单规则以难以预测的方式产生出复杂行为，这种系统的宏观行为有时也称为涌现（emergent）。这样就有了复杂系统的另一个定义：具有涌现和自组织行为的系统。复杂性科学的核心问题是：涌现和自组织行为是如何产生的。在书中我会尝试从各种角度来阐释这个问题。

如何度量复杂性

前面我介绍了复杂系统的一些性质。但是还有量的问题：一个特定的复杂系统到底有多复杂呢？也就是说，我们该如何度量复杂性？可以精确地说出一个系统比另一个复杂多少吗？这个问题很重要，但是还没有完全解决，至今仍是充满争议的领域。在第7章我们会看到，有许多度量复杂性的方式；不过还没有哪一种得到公认。书中许多章节描述了复杂性的各种度量方法及其用途。

但是如果公认的复杂性定义都没有，又如何会有复杂性科学呢？

对这个问题我有两个回答。首先，虽然有很多书和文章使用这些术语，但是既不存在单独的复杂性科学，也不存在单独的复杂性理论。其次，我在书中会反复提到，一门新的科学形成的过程，就是不断尝试对其中心概念进行定义的过程。对信息、计算、序和生命等核心概念的定义就是这样的例子。书中我会对这些奋斗历程的历史和现状进行阐述，并将它们与我们对复杂性的理解结合起来。这本书讲的是科学前沿，但也讲述科技前沿背后的核心概念的历史，下面四章讲的就是贯穿全书的核心概念的历史和背景。

第2章　动力学、混沌和预测

再一次一无所知[14]，从头开始……这让我很开心。

对于我们来说平常大小的事物，人们为之写诗的那些——云彩、水仙花、瀑布，它们对于我们，就好像天堂对于古希腊人，充满了神奇……现在也许是最好的时代，你曾以为正确的东西几乎都是错的。

——斯托帕德（Tom Stoppard），

《阿卡狄亚》（Arcadia ）

动力系统理论（动力学，dynamics）关注的是对系统的描述和预测，其所关注的系统通过许多相互作用的组分的集体行为涌现出宏观层面的复杂变化。动力一词意味着变化。而动力系统则是以某种方式随时间变化的系统。下面是动力系统的一些例子：

◆太阳系（行星位置随时间变化）；

◆心脏（周期性跳动）；

◆大脑（神经元不断激发，神经递质在神经元之间传递，突触强度变化，整个系统不断变化）；

◆股票市场；

◆世界人口；

◆全球气候。

不仅这些，其他你想得到的系统几乎都是动力系统。甚至岩石在地理时间尺度上也是变化的。动力系统理论以最一般化的方式描述系统的变化，描述变化可能的宏观形态，以及对于其变化能够做出怎样的估计和预测。

近年来，动力系统理论很受大众关注，这是因为它的一个分支——混沌学——发现了一些让人着迷的结果。但实际上它的历史很悠久，同许多学科一样，它可以追溯到古希腊哲学家亚里士多德。

动力系统理论的起源

亚里士多德（图2.1）是目前所知的最早论述运动理论的人之一，他的理论流行了1500多年。他的理论有两个主要原理，后来发现都是错的。首先，他认为地面上的运动与天上的不同。他认为地面上的物体在受到力推动时才会沿直线运动；没有力，物体就会保持静止。而在天上，行星等天体是围绕着地球不断做圆周运动。另外，亚里士多德认为，在地面上，不同物质组成的物体运动方式也不一样。比如，他认为石头落向地面是因为石头主要是由土元素组成，而烟会上升则是因为烟是由气元素组成。在天上也是一样，越重的物体中的土元素越多，下落也越快。

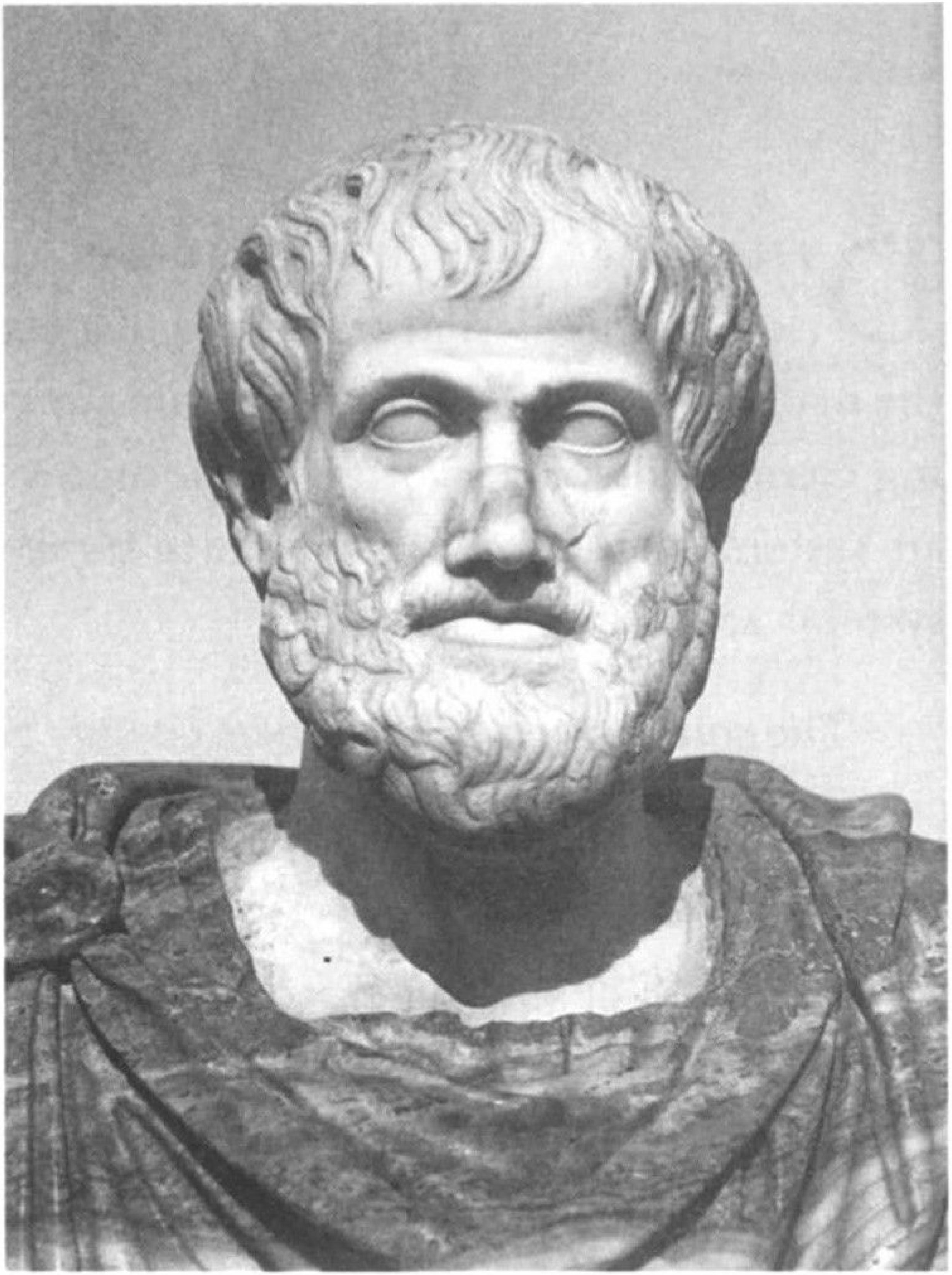


图2.1　亚里士多德（前384—前322）（卢多威斯收藏）（Ludovisi Collection）

同以前许多理论家一样，亚里士多德在构造理论时没有考虑实验验证。他的方法是用逻辑和常识引导理论；用实验对理论进行验证的重要性在当时还没有被认识到；亚里士多德的思想影响很大，一直统治着西方科学，直到16世纪——伽利略（图2.2）登上历史舞台。

伽利略、他之前的哥白尼以及与他同时代的开普勒是实验和观察科学的先驱。哥白尼提出行星不是围绕地球而是围绕太阳运行。（伽利略在宣扬这种观点时受到了天主教会的强烈阻挠，最后被迫公开宣布放弃。直到1992年教会才正式承认对伽利略的迫害是错误的。）在16世纪初，开普勒发现行星的运行轨迹不是圆而是椭圆，他还发现了关于这种椭圆运动的几条定律。

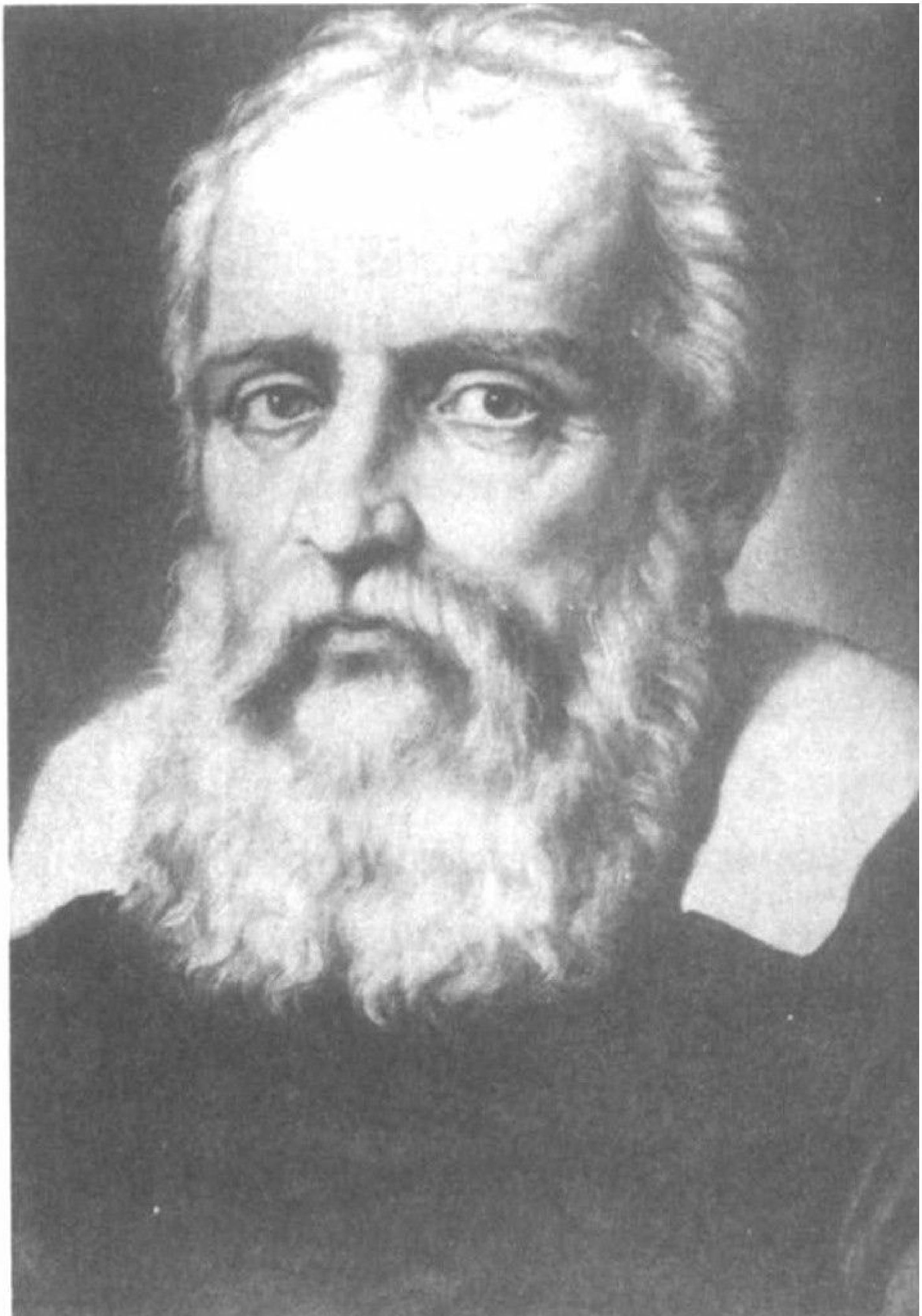


图2.2　伽利略（1564—1642）（美国物理学会西格尔图像档案，斯科特·贝尔收藏）

哥白尼和开普勒只研究了天体的运动，而伽利略不仅研究天上的运动，也研究地面上的，他做了一些我们现在在中学物理课上会学到的实验：单摆、沿斜面滚动的小球、自由落体、镜面光线反射。不过伽利略可没有我们现在使用的那些精密实验设备，据说他通过数脉搏来计算单摆的摆动周期，还在比萨斜塔上下落物体以测量重力的效应。这些经典实验彻底改变了对运动的理解，并且直接驳斥了长期盛行的亚里士多德的观点。与直觉不同，静止并不是物体的自然状态；相反，要施加力才能让运动物体停下来。不管物体多重，在真空中下落的速度都是一样的。最具革命性的是，地面上的运动定律居然也能解释天上的运动。自从伽利略之后，有了实验观察作为基础，科学革命的发生就不可避免了。

动力学历史上最重要的人物是牛顿（图2.3），牛顿生于伽利略死后那一年。他可以说是凭一己之力创建了动力学。为了创建动力学，他还要先发明微积分——描述运动和变化的数学。



图2.3　牛顿（1643—1727）（不知名艺术家雕刻，由美国物理学会西格尔图像档案提供）

物理学家将对运动的总体研究称为机械力学（mechanics）。这个词源自古希腊，因为古典观点认为，所有运动都可以用杠杆、滑轮、轮轴等简单“机械”的动作组合来解释。牛顿的工作现在被称为经典力学。力学分为两部分：描述物体如何运动的运动学（kinematic），以及解释

物体为何遵循运动学定律的动力学。例如开普勒定律就是运动学定律，它们描述了行星如何运动（以太阳为焦点沿椭圆运动），但没有解释行星为何这样运动。牛顿的定律则是动力学的基础，它们用力和质量作为基本概念解释了一切物体的运动，包括行星。

下面是著名的牛顿三大定律：

1.在任何情况下，一切物体在不受外力作用时，总保持静止或匀速直线运动状态。

2.物体的加速度与物体的质量成反比。

3.两个物体之间的作用力和反作用力，在同一条直线上，大小相等，方向相反。

牛顿的伟大之处在于他认识到这些定律不仅适用于地面上的物体，对天上的物体也同样适用。匀速运动定律是伽利略首先提出来的，但是他认为只适用于地面上的物体。而牛顿则认为这条定律对行星应该也适用，并且认识到需要用力（引力）来解释椭圆运动方向的不断变化。牛顿的另一重要贡献是提出了万有引力定律：两个物体之间的引力与两者质量的乘积成正比，与两者距离的平方成反比。牛顿深刻认识到这条定律适用于宇宙中一切事物，无论是行星还是苹果，这个认识是现代科学的基石。正如他说的："自然简单而自足，[15] 对宏大物体的运动成立的，对微小物体也同样成立。"

牛顿力学描绘了一幅"钟表宇宙"的图景：设定好初始状态，然后就遵循着三条定律一直运行下去。数学家拉普拉斯认识到其中蕴含了可以如钟表般精准预测的观念：他在1814年断言，根据牛顿定律，只要知道宇宙中所有粒子的当前位置和速度，原则上就有可能预测任何时刻的情况。[16] 在20世纪40年代计算机被发明出来之后，这种"原则上"的可能似乎有可能变成现实了。

对预测的重新认识

然而，20世纪的两个重要发现表明，拉普拉斯的精确预测的梦想，即使在原则上也是不可能的。1927年，海森堡（Werner Heisenberg）提

出了量子力学中的“测不准原理”，证明不可能在准确测量粒子位置的同时，又准确测量其动量（质量乘以速度）。对于其位置知道得越多，对于其动量就知道得越少，反过来也是一样。不过，海森堡原理还只是限制了对量子世界微观粒子的测量，大多数人都只是觉得它挺有趣，但是对宏观尺度上的预测——比如天气预报——应该没有多大影响。

然而混沌的发现给了精确预测的梦想最后一击。混沌指的是一些系统——混沌系统——对于其初始位置和动量的测量如果有极其微小的不精确，也会导致对其的长期预测产生巨大的误差。也就是常说的“对初始条件的敏感依赖性”。

对于一些自然系统，并没有这个问题。如果你对初始条件的测量不是十分精确，你的预测即使不全对，也会八九不离十。例如天文学家在测量行星位置时即使误差较大，也还是能准确预测日食。而对初始条件的敏感依赖性指的是，如果系统是混沌的，在测量初始位置时即使只有极其微小的误差，在预测其未来的运动时也会产生巨大的误差。对于这样的系统（飓风就是例子），一点点误差，不管多小，也会导致长期预测很不精确。

这一点很不符合直觉，事实上，很长一段时间里，科学家们都认为这不可能。然而，混沌现象在很多系统中都被观测到了，心脏紊乱、湍流、电路、水滴，还有许多其他看似无关的现象。现在混沌系统的存在已成为科学中公认的事实。

现在已无法说清楚是谁最先意识到可能存在这类系统。远在量子力学出现之前，就有很多人提出了对初始条件敏感依赖性的可能性。例如，物理学家麦克斯韦（James Clerk Maxwell）在1873年就猜想，有些量的“物理尺度太小，[17] 以致无法被有局限性的人类注意，却有可能导致极为重要的结果”。

第一个明确的混沌系统的例子可能是19世纪末由法国数学家庞加莱（Henri Poincaré）（图2.4）给出。庞加莱是现代动力系统理论的奠基者，可能也是贡献最大的人，大力推动了牛顿力学的发展。庞加莱在试图解决一个比预测飓风简单得多的问题时发现了对初始条件的敏感依赖性。他试图解决的是所谓的三体问题（three-body problem）：用牛顿定律预测通过引力相互作用的三个物体的长期运动。牛顿已经解决了二体

问题。但没想到三体问题要复杂得多。在向瑞典国王表示敬意的一次数学竞赛中，庞加莱将其解决了。竞赛主办方提供2500瑞典克朗奖励解决“多体”问题：用牛顿定律预测任意多个相互吸引的物体的未来运动。提出这个问题是为了确定太阳系是否稳定，行星是会维持还是会偏离目前的轨道？庞加莱想先试着解决三体问题。

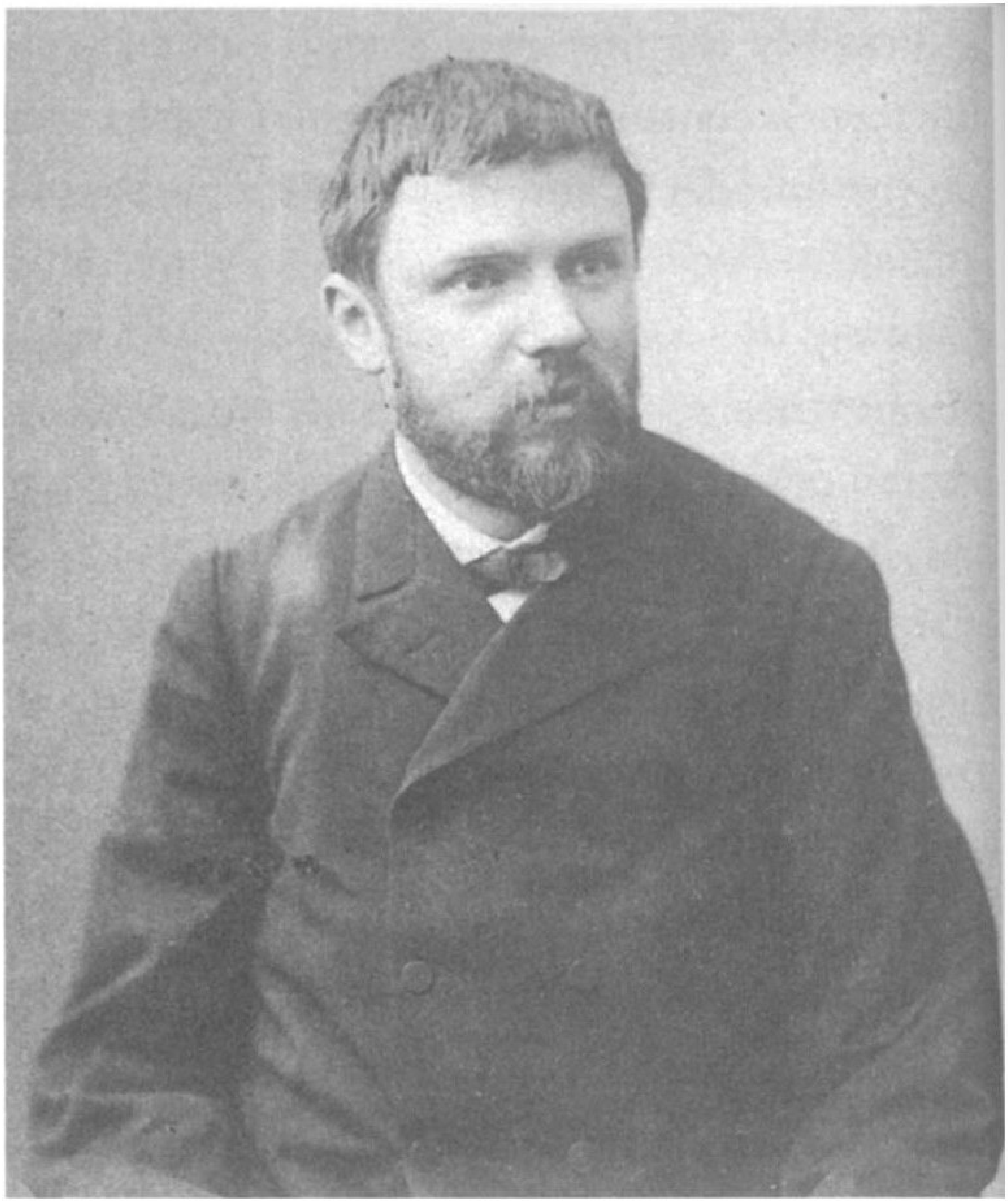


图2.4　庞加莱（1854—1912）（美国物理学会西格尔图像档案）

他并没有完全成功——这个问题实在太复杂了。但是他的尝试很精彩，所以最后还是赢得了奖金。牛顿发明了微积分，而庞加莱为了解决这个问题也创建了一个新的数学分支——代数拓扑（algebraic

topology）。拓扑学是几何学的扩展，正是在研究三体问题的几何结果的过程中，庞加莱发现了对初始条件的敏感依赖性。下面是他对此的总结：

> 如果我们能知道自然界的定律1和宇宙在初始时刻的精确位置，我们就能精确预测宇宙在此后的情况。但是即便我们弄清了自然界的定律，我们也还是只能近似地知道初始状态。如果我们能同样近似地预测以后的状态，这也够了，我们也就能说现象是可以预测的，而且受到定律的约束。但并不总是这样，初始条件的细微差别有可能会导致最终现象的极大不同。前者的微小误差会导致后者的巨大误差。预测变得不可能……

换句话说，即便我们完全知道了运动定律，两组不同的初始条件（在这里是指物体的初始位置、质量和速度），即使差别很小，有时候也会导致系统随后的运动极为不同。庞加莱在三体问题中发现了一个这样的例子。

直到电子计算机出现之后，科学界才开始认识这类现象的意义。庞加莱远远超越了他所处的时代，他意识到对初始条件的敏感依赖性将会阻碍对天气的长期预报。他的远见于1963年被证实，气象学家洛伦兹（Edward Lorenz）发现，[18] 即使是很简单的计算机气象模型，也会有对初始条件的敏感依赖性。现在虽然有了高度复杂的气象计算模型，天气预报也最多只能做到大致准确预测一个星期。目前还不清楚这个局限是否是天气的混沌本质导致的，也不知道通过收集更多数据和构造更好的模型，可以将这个局限推进多远。

线性兔子和非线性兔子

现在我们再详细了解一下对初始条件的敏感依赖性。混沌系统中初始的不确定性到底是如何被急剧放大的呢？关键因素是非线性。对于线性系统，你可以先了解其组成，然后将它们合到一起。当我的两个儿子和我一起做厨艺时，他们喜欢轮流加原料。杰克放两杯面粉，跟着尼克又加一杯糖。结果呢？三杯面粉和糖的混合物，整体等于部分之和。

对于非线性系统，整体则不等于部分之和。杰克放了两杯苏打粉，尼克又加了一杯醋。整个事情就不可收拾了（你可以自己在家里试

试）。有什么后果？你会得到大量醋、苏打粉和二氧化碳混合的泡泡。两者之间的区别在于：前面的糖和面粉不会产生反应生成新的东西，而后者的醋和苏打粉会剧烈反应，产生很多二氧化碳。

还原论者喜欢线性，而非线性则是还原论者的梦魇。理解线性和非线性的区别很有用，值得研究一下。为了更好地理解非线性以及混沌现象，我们要研究一点点简单的数学，借用一个经典的生物群体数量动力学模型来阐释线性和非线性。设想你养了一群兔子，兔子会配对生小兔子，每对兔子父母每年会生4只小兔子然后死去。图2.5显示了兔子的繁殖状况。

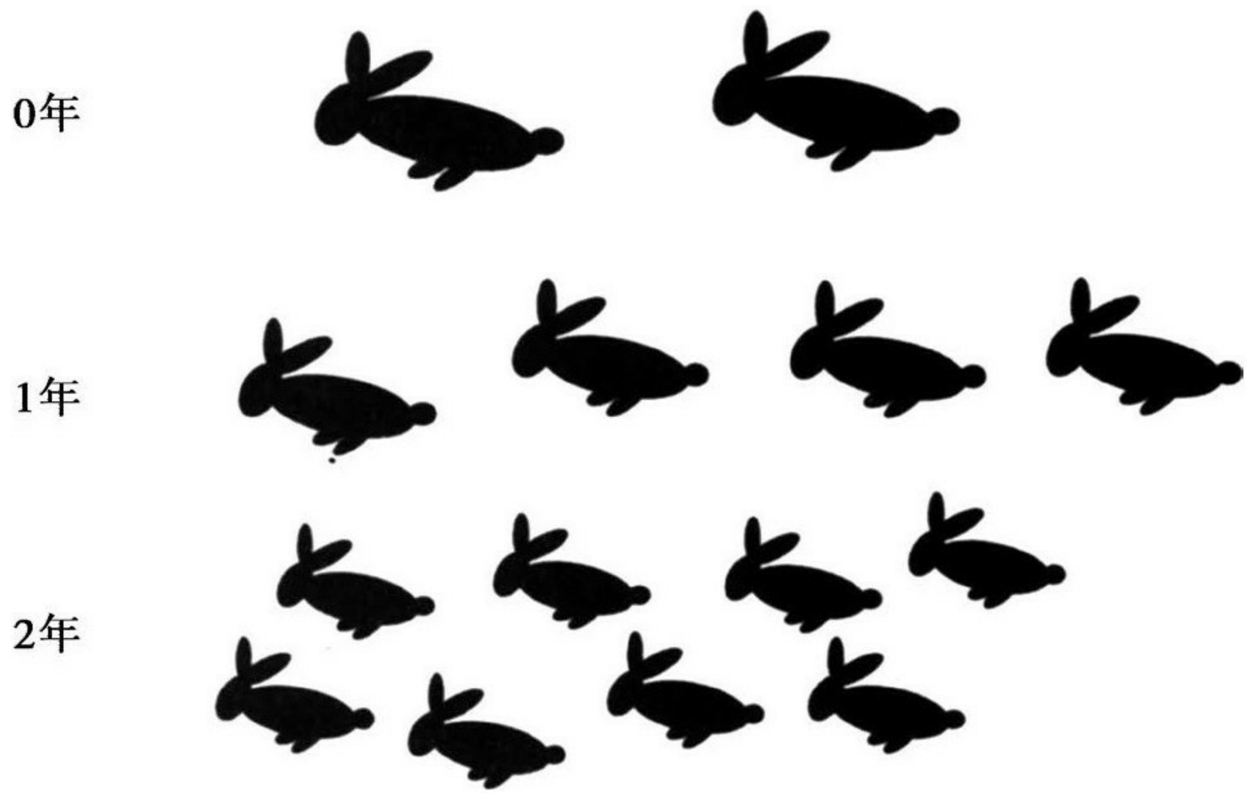


图2.5　倍增的兔群

很显然，如果不受限制，兔子的数量会每年翻番（这意味着兔子很快会接管这个星球，乃至太阳系和整个宇宙，不过我们暂时还不用担心）。

这是一个线性系统：[19] 整体等于部分之和。我想让它们做什么呢？我们先将4只兔子分开放到两个岛上，每个岛上2只。然后让兔子继

续繁殖。图2.6显示了繁殖两年的情形。

两边都是每年翻番。不管是哪一年，如果你把两个岛的兔子加起来，你得到的数量还是与没分开时一样多。

如果以当年的兔子数量为横坐标，以次年的兔子数量为纵坐标，将各年的数据标上去，你将会得到一条直线（图2.7）。这就是为什么称之为线性系统。

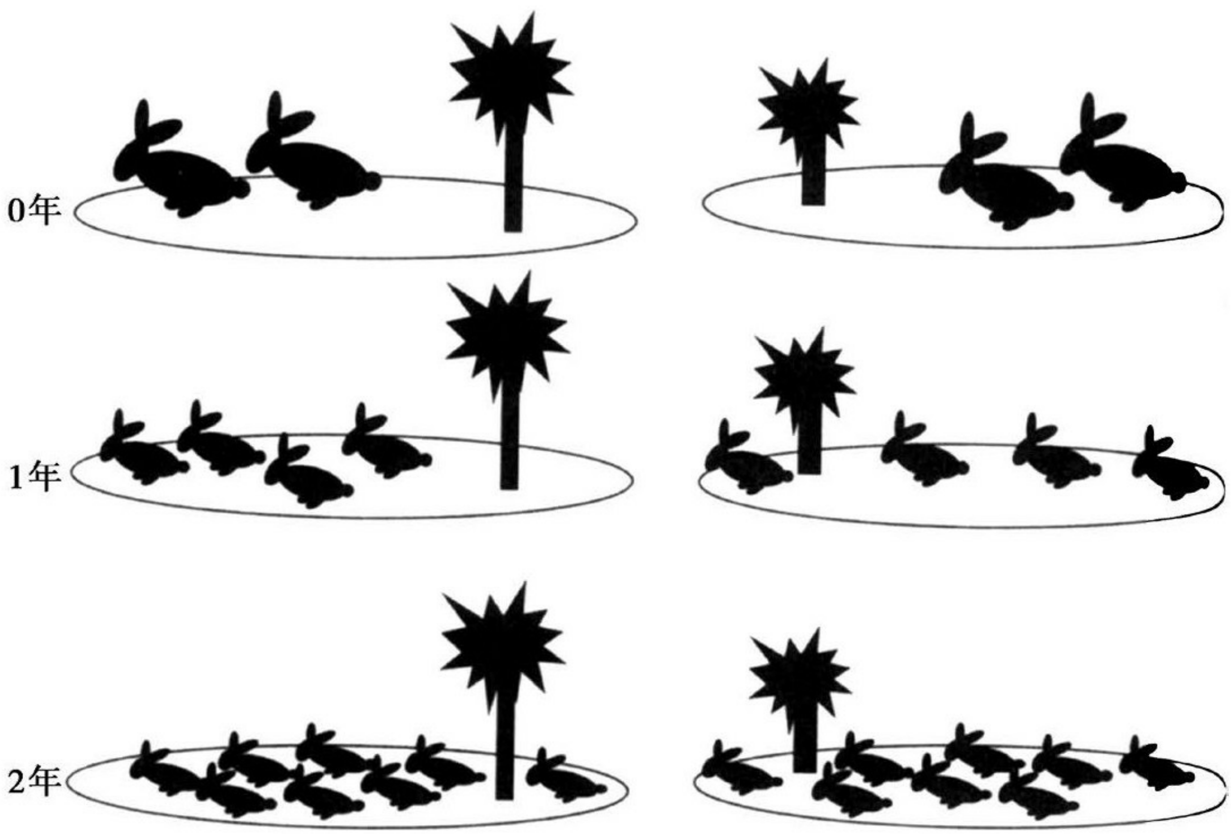


图2.6　倍增的兔群，分开在两个岛上

但是如果考虑到种群数量增长所受的限制，情况会怎样呢？这会使得增长规则变为非线性的。假定前面的规则仍然成立，每对兔子每年生4只小兔子然后死去。不过现在有些小兔子会因为太过拥挤没有繁殖就死去。研究种群数量的生物学家常用逻辑斯蒂模型[20]（Logistic model）描述这种情形下群体数量的增长。这个模型以一种简化方式描述群体数量的增长。你设定好出生率、死亡率（由于种群数量过多导致的死亡概率）以及最大种群承载能力（栖息地所能承载的种群数量上限），然后将这一代的种群数量代入逻辑斯蒂模型，就能算出下一代的

种群数量。在这里我不给出逻辑斯蒂模型的具体形式[21]（注释中有），你可以在图2.8中看到它的变化情况。

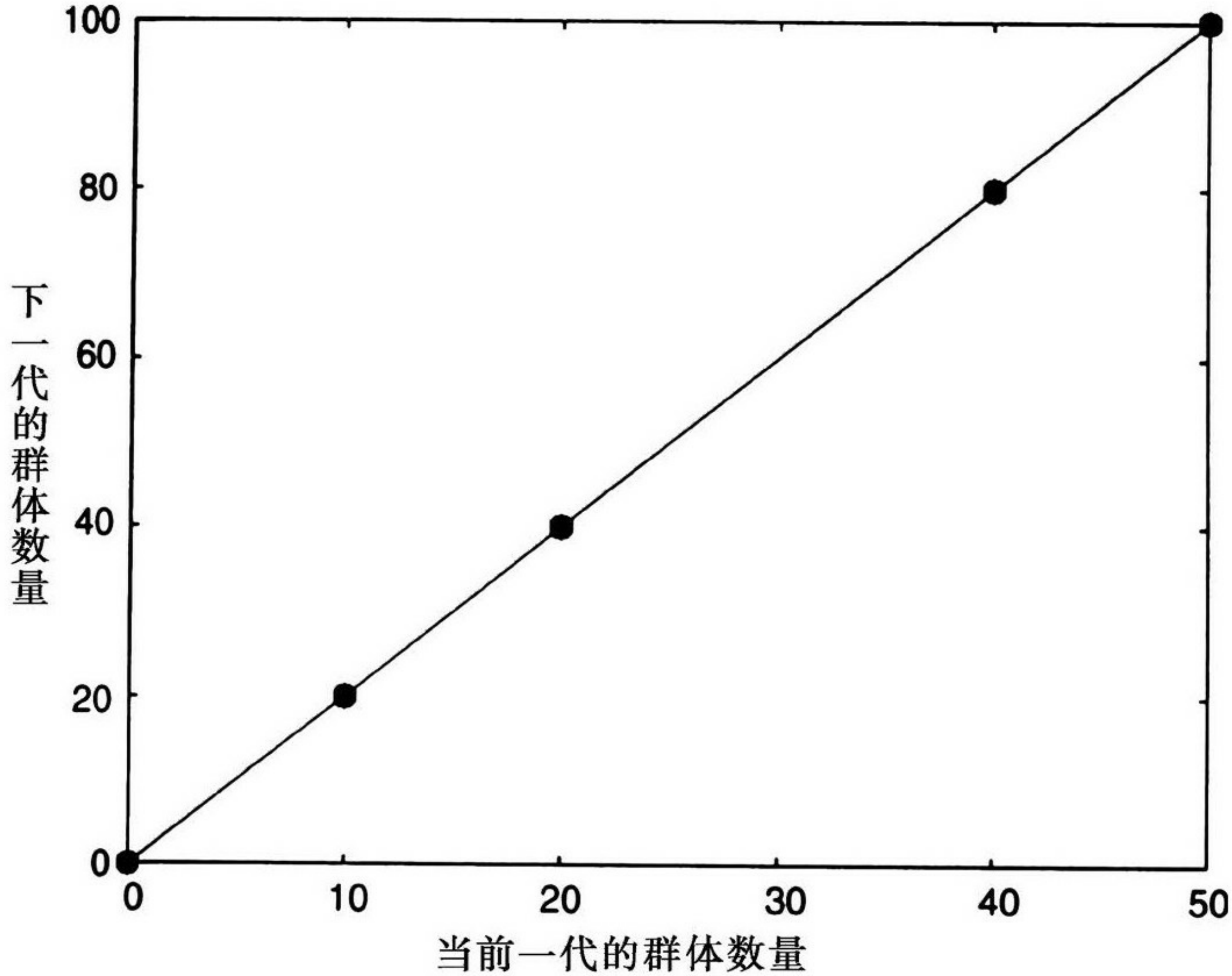


图2.7 线性模型中当年与次年种群数量的关系曲线

举个简单的例子，设出生率为2，死亡率为0.4，承载力为32，第一代有20只兔子。用逻辑斯蒂模型算出第二代为12只。将新的种群数量再代进去，又可以得出第三代仍然是12只兔子存活。此后的兔子数量将一直维持在12只。

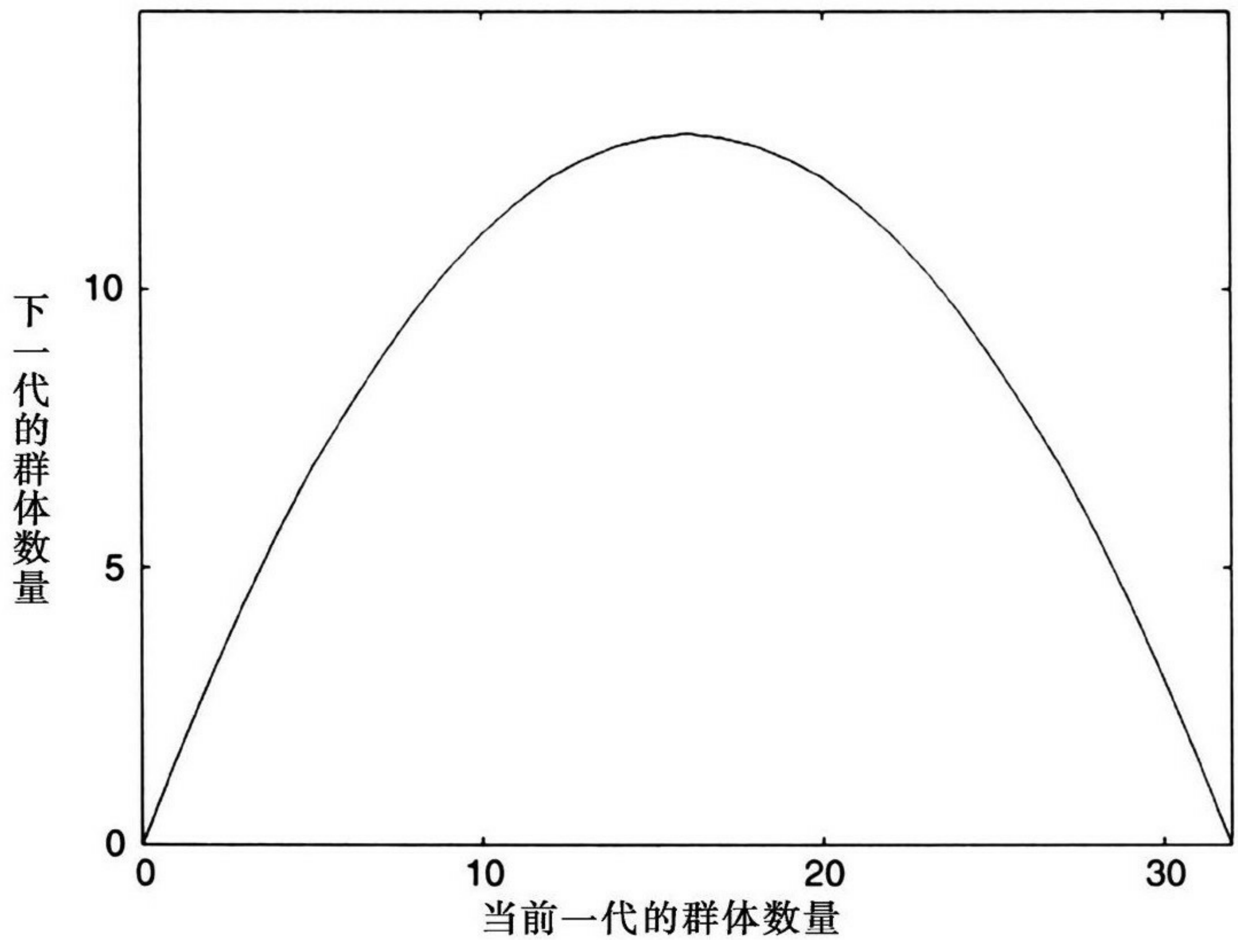


图2.8　根据逻辑斯蒂模型得出的当年与次年种群数量的关系曲线，出生率为2，死亡率为0.4，承载力为32。如果取其他参数，曲线仍然是抛物线

如果将死亡率降到0.1（其他参数不变），会有些有趣的事情发生。根据模型可以得出第二代为14.25只兔子，第三代则为15.01816只。

等一下！怎么会有0.25只兔子，还有稀奇古怪的0.01816只？在真实世界中显然是不可能的，不过这只是模型，允许兔子数量为小数。这样在数学上简单些，而且预测的兔子数量仍然大致符合实际。所以这里我们无须为此担心。

将算出的种群数量再代进去计算下一代的种群数量，这个不断重复的过程即所谓的“对模型进行迭代”。

如果将死亡率恢复成0.4，承载力翻一倍变成64，结果又会怎样呢？根据模型我们发现，从20只兔子出发，9年后种群数量会变为接近

24的一个值，然后停在那里。

你可能注意到了这些例子中的种群变化比前面单纯每年翻番的情形复杂得多。这是因为引入了种群数量过多导致的死亡，模型变成了非线性的。其图形不再是直线，而是抛物线（图2.8）。逻辑斯蒂模型中的群体数量变化不再简单等于部分之和。为了说明这一点，我们将20只兔子分为两群，每群10只，再对各群进行迭代（参数同前面一样，出生率为2，死亡率为0.4）。图2.9为迭代结果。

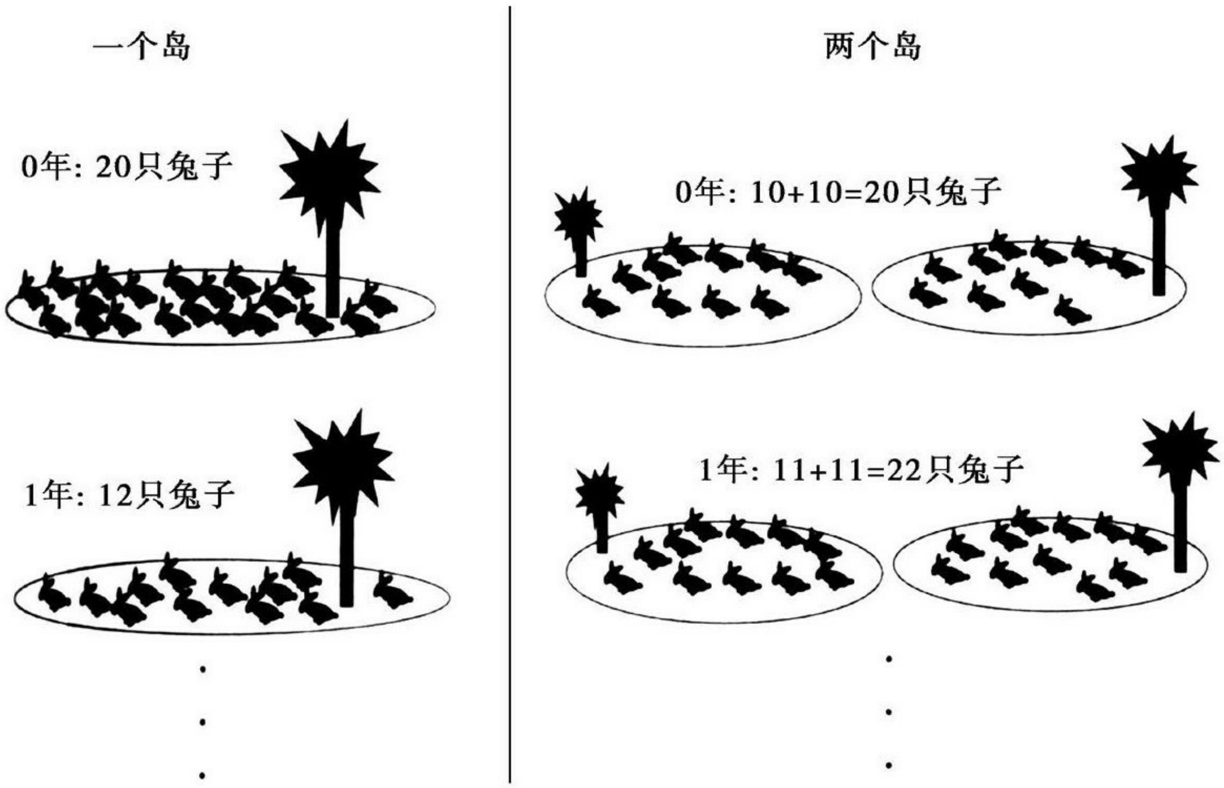


图2.9　分到两个岛上的兔子，逻辑斯蒂模型

第一年，前面是20只兔子只剩下12只，而分成两群后，每群有11只，总共22只。整体的变化不再等于各部分的变化之和。

逻辑斯蒂映射[22]

许多研究这一类事物的科学家和数学家使用逻辑斯蒂模型的一个简化形式，逻辑斯蒂映射（logistic map），它也许是动力系统理论和混沌研究中最著名的方程。逻辑斯蒂映射中出生率和死亡率的效应被合成一

个数，记作R。种群规模用“承载率”替代，记为x。这个简化模型问世之后，科学界和数学界很快就将种群规模、承载力等与现实世界的联系抛到脑后，转而着迷于这个方程本身，因为它的特性太让人震惊了。现在我们也来体验一下。

下面就是这个方程，其中x_t 是当前值，x_{t+1} 则是下一步的值：[23]

$$x_{t+1}=Rx_t\ (1-x_t)$$

我给出逻辑斯蒂映射的方程是为了向你展示它有多简单。事实上，它是能抓住混沌本质——对初始条件的敏感依赖性——的最简单的系统之一。1971年，数学生物学家梅（Robert May）在著名的《自然》杂志上发表了一篇文章 [24] 分析逻辑斯蒂映射，引起了种群生物学家的关注。在此之前也有一些数学家对其进行了详细分析，包括乌拉姆（Stanislaw Ulam）、冯·诺依曼（John von Neumann）、梅特罗波利斯（Nicholas Metropolis）、保罗·斯坦（Paul Stein）和米隆·斯坦（Myron Stein）。[25] 但它真正变得有名是在20世纪80年代，物理学家费根鲍姆（Mitchell Feigenbaum）利用它展示了一大类混沌系统的共性。由于其显然的简单性和深厚的历史，它成了介绍动力系统理论和混沌的一些主要概念的完美载体。

如果我们让R的值变化，逻辑斯蒂映射就变得非常有趣。我们先从R=2开始。x的初始值x_0 也必须介于0和1之间，姑且设为0.5。将它们代入逻辑斯蒂映射，得出x_1 为0.5。同样，x_2 也是0.5，后面也一样。因此，如果R=2，种群初始值为最大值的一半，以后就会一直不变。

现在让x_0 =0.2。你可以自己用计算器算一下（我用的一个最多显示7位小数的计算器）。结果更有意思了：

x_0 =0. 2

x_1 =0. 32

x_2 =0. 4352

x_3 =0. 4916019

x_4 =0. 4998589

x_5 =0. 5

x_6 =0. 5

……

最终结果是一样的（永远是x_t =0.5），但是迭代了5次才得到。

用图可以看得更清楚。图2.10是x_t 在前20步的值的图形。我用线将这些点连起来了，这样可以更清楚地看到，随着时间推移，x迅速收敛到0.5。

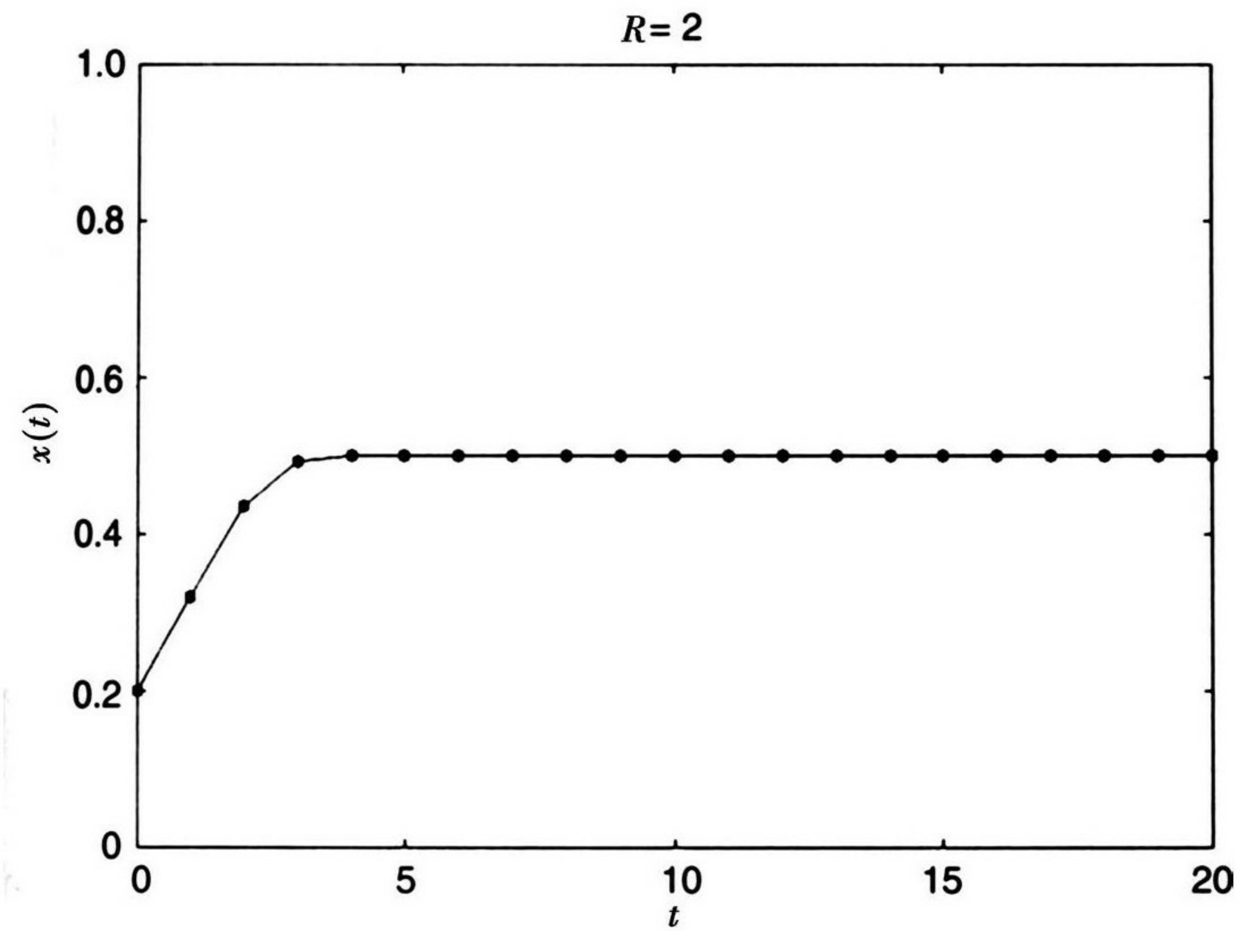


图2.10　R=2，x_0 =0.2时逻辑斯蒂映射的变化情况

如果x_0 很大，比如0.99，又会怎样呢？图2.11显示了得到的图形。

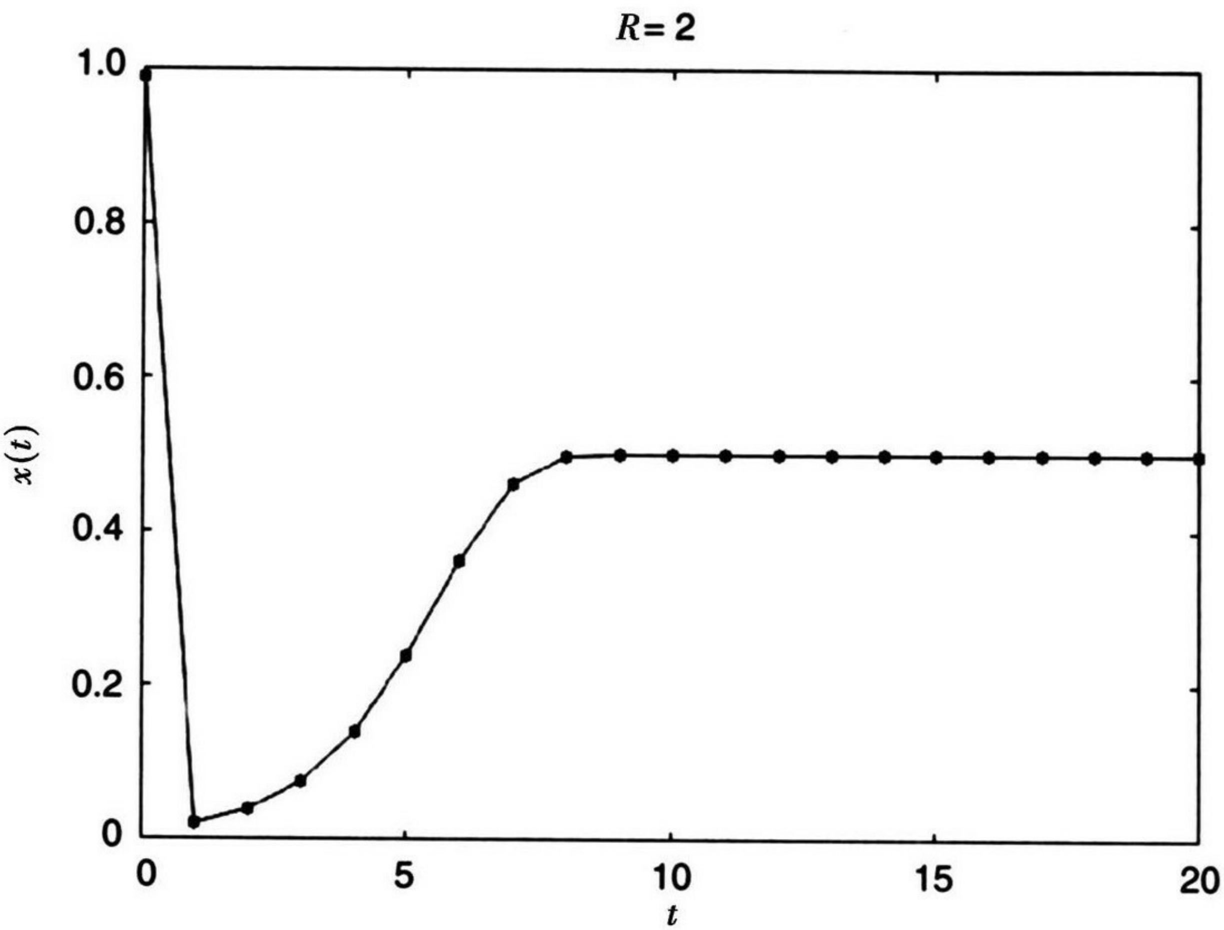


图2.11　R=2，x_0 =0.99时逻辑斯蒂映射的变化情况

最终的结果还是一样的，不过过程要长一些，波动也更剧烈。

你可能已经猜到了：只要R=2，x_t 最终都会到达0.5，并停在那里。0.5正是所谓的不动点（fixed point）：到达这一点所花的时间依赖于出发点，但是一旦你到达了那里，你就会保持不动。如果你愿意，可以让R=2.5，再试一下，同样你会发现系统总是到达一个不动点，不过这次不动点是0.6。

R=3.　1的情形更有趣。逻辑斯蒂映射的变化更加复杂了。图2.12是x_0 =0.2时的图形。

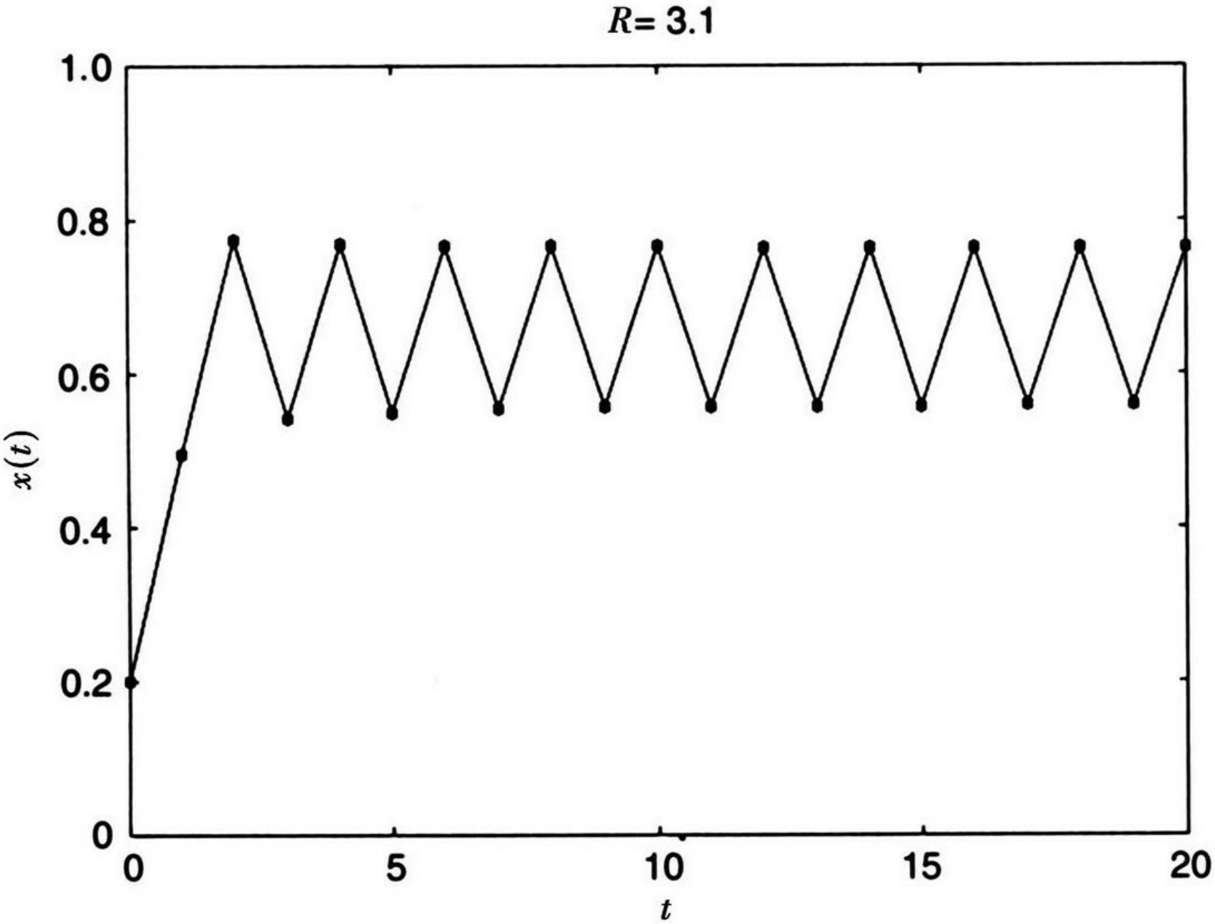


图2.12　R=3.1，x_0 =0.2时逻辑斯蒂映射的变化情况

在这个例子中，x永远也不会停在一个不动点；它最终会在两个值（0.5580141和0.7645665）之间振荡。如果将前者代入方程，就会得到后者，反过来也是一样，因此振荡会一直持续下去。不管x_0 取什么值，最后都会形成这个振荡。这种最终的变化位置（无论是不动点还是振荡）被称为“吸引子”，这个说法很形象，因为任何初始位置最终都会“被吸引到其中”。

往上一直到R等于大约3.4，逻辑斯蒂映射都会有类似的变化：在迭代一些步骤后，系统会在两个不同的值之间周期振荡（最终的振荡点由R决定）。因为是在两个值之间振荡，系统的周期为2。

但是如果R介于3.4和3.5之间，情况又突然变了。不管x_0　取何值，系统最终都会形成在四个值之间的周期振荡，而不是两个。例如，如果R=3.49，x_0 =0.2，最终的结果就像图2.13那样。

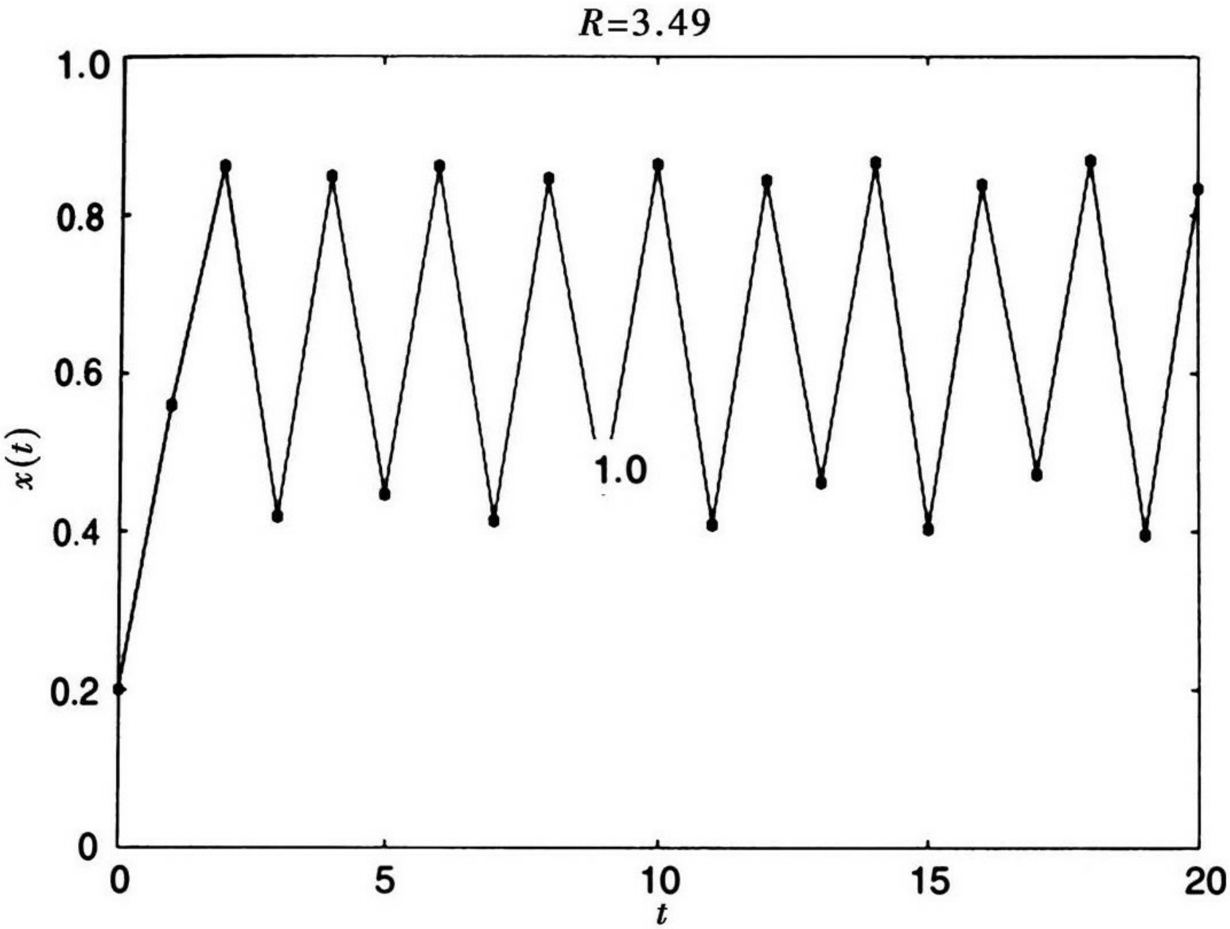


图2.13　R=3.49，x_0 =0.2时逻辑斯蒂映射的变化情况

x的值很快就开始在四个不同的值之间周期振荡（如果你想知道，它们分别大约是0.872，0.389，0.829和0.494）。也就是说，在3.4和3.5之间的某个R值，最终的振荡周期突然从2增到4。

在3.54和3.55之间的某个R值，周期再次突然倍增，一下跃升到8。在3.564和3.565之间的某个值周期跃升到16。在3.5687和3.5688之间周期又跃升到32。周期一次又一次倍增，前后R的间隔也越来越小，很快，在R大约等于3.569946时，周期已趋向于无穷。在此之前，逻辑斯蒂映射的变化大致都可以预测。如果R值给定，从任何x_0 点出发的最终长期变化都能预测得到：R小于3.1时会到达不动点，R介于3.1和3.4之间时会形成双周期振荡，等等。

但是当R等于大约3.569946时，x的值不再进入振荡，它们会变成混

沌。[26] 下面解释一下。将x_0 ，x_1 ，x_2 ……的值组成的序列称为x的轨道。在产生混沌的R值，让两条轨道从非常接近的x_0 值出发，结果不会收敛到同一个不动点或周期振荡，相反它们会逐渐发散开。在R=3.569946时，发散还很慢，但如果将R设为4.0，我们就会发现轨道极为敏感地依赖于x_0 。我们先将x_0 设为0.2，对逻辑斯蒂映射进行迭代，得到一条轨道。然后细微地变动一下x_0 ，让x_0 =0.2000000001，再对逻辑斯蒂映射进行迭代，得到第二条轨道。图2.14中的实心圆圈连成的实线就是第一条轨道，空心圆圈连成的虚线则是第二条轨道。

这两条轨道开始的时候很接近（非常接近，以至于实线轨道把虚线轨道都盖住了），但在大约30次迭代之后，它们明显分开了，很快就不再具有相关性。这就是“对初始条件的敏感依赖性”的由来。

我们已经看到有三种不同的最终状态（吸引子）：不动点、周期和混沌（混沌吸引子有时候也称为“奇怪吸引子”）。吸引子的类型是动力系统理论刻画系统行为的一种方式。

我们再仔细来看看混沌行为到底有多不寻常。逻辑斯蒂映射极为简单，并且完全是确定性的：每个x_t 值都有且仅有一个映射值x_{t+1} 。然而得到的混沌轨道看上去却非常随机——事实上逻辑斯蒂映射还被用来在计算机中生成伪随机数 [27] 。因此，表面上的随机可以来自非常简单的确定性系统。

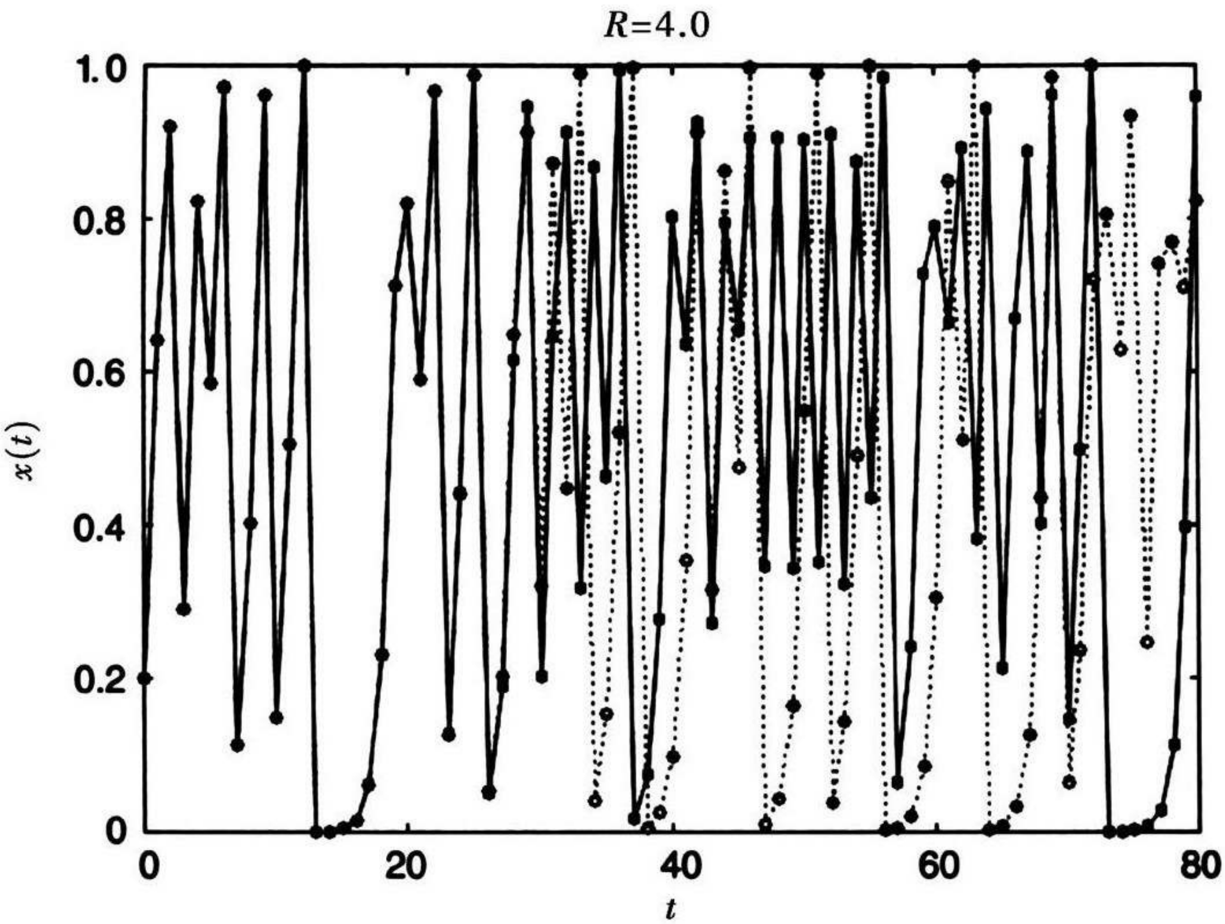


图2.14　R=4.0时逻辑斯蒂映射的两条轨道：x_0 =0.2和x_0 =0.2000000001

此外，对于产生混沌的R值，如果初始条件x_0 有任何的不确定性，对一定时间之后的轨道就无法再预测了。R=4时我们已经看到这种状况。如果我们对x_0 不能精确到小数点后第10位——大多数实验观察都做不到这么精确——那么大约在t=30时，x_t 的值就无法预测了。对于任何能产生混沌的R值，只要x_0 有不确定性，不管精确到小数点后多少位，最终都会在t大于某个值时变得无法预测。

数学生物学家梅对这些惊人的特性进行了总结，与庞加莱遥相呼应：

简单的确定性方程 [28] （1）（即逻辑斯蒂映射）能产生类似于随机噪声的确定性轨道，这个事实有着让人困扰的实际含义。例如，这就意味着种群调查数据中那种明显的不稳定波动不一定表明环境的变

化莫测或是采样有错误：它们有可能就是像方程（1）这样完全确定性的种群数量变化关系所导致的……另外，还可以看到，在混沌中，不管初始条件有多接近，在足够长的时间之后，它们的轨道还是会相互分开。这意味着，即使我们的模型很简单，所有的参数也都完全确定，长期预测也仍然是不可能的。

简而言之，系统存在混沌也就意味着，拉普拉斯式的完美预测不仅在实践中无法做到，在原则上也是不可能的，因为我们永远也无法知道 x_0 小数点后的无穷多位数值。这是一个非常深刻的负面结论，它与量子力学一起，摧毁了19世纪以来的乐观心态——认为牛顿式宇宙就像钟表一样沿着可预测的路径运行。

但是对逻辑斯蒂映射的研究是不是也会产生一些正面作用呢？对于试图发现随时间变化的系统的一般原则的动力系统理论，它能有所助益吗？事实上，对逻辑斯蒂等映射的深入研究也已经得到了同样深刻的正面结果——从中发现了混沌系统的普遍特征。

混沌的共性

最早用术语混沌来描述对初始条件具有敏感依赖性的动力系统的人是物理学家李天岩（T.Y.Li）和约克（James Yorke）。[29] 这个词用得恰到好处：在口语中“混沌”一词意指随机和不可预测，在逻辑斯蒂映射的混沌中就有这些性质。然而，与口语中的混沌不同，数学混沌还有本质上的秩序，即很多混沌系统所共有的普适性。

第一条普适性质：通往混沌的倍周期之路

在前面的数学探讨中，我们看到随着R从2.0增大到4.0，逻辑斯蒂迭代最初会产生不动点，然后是2周期振荡，然后是4周期，然后是8周期，一直下去，直到出现混沌。在动力系统理论中，这些突然的周期倍增被称为分叉（bifurcation）。不断分叉直至混沌的过程就是“通往混沌的倍周期之路”。

我们经常用分叉图来表现分叉，分叉图是“控制参数”（比如R）和系统吸引子之间的函数关系。图2.15就是逻辑斯蒂映射的分叉图。横坐标为R，纵坐标是各R值对应的x的最终值（吸引子）。例如，R=2.9

时，x会到达固定点吸引子x=0.655。R=3.0时，x会到达双周期吸引子。这就是图中第一个分叉点，不动点吸引子换成了双周期吸引子。在3.4和3.5之间，又分叉为4周期吸引子，后面不断周期倍增，直至R到达3.569946附近，开始出现混沌的发端（onset of chaos）。

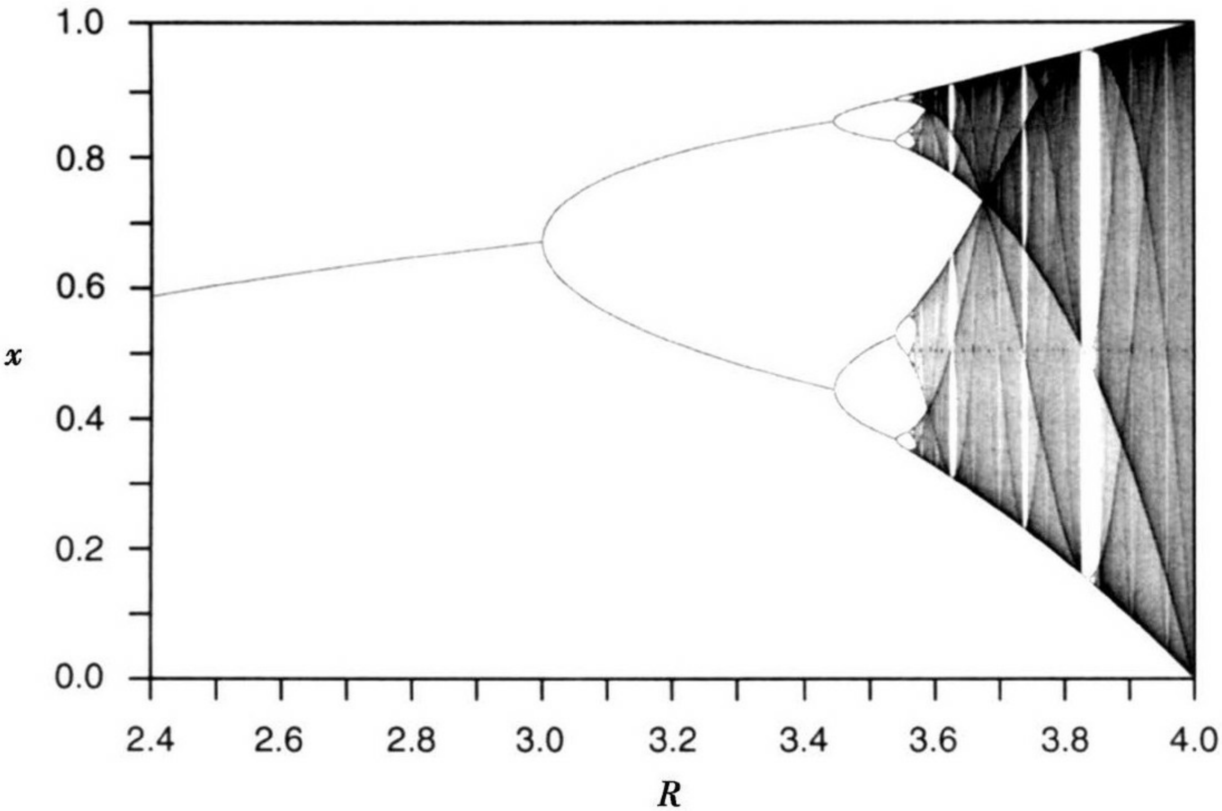


图2.15　逻辑斯蒂映射分叉图，用吸引子作为R的函数

通往混沌的倍周期之路有着悠久的历史。[30] 早在20世纪20年代，就在数学方程中发现了倍周期分叉，20世纪50年代芬兰数学家米尔堡（P.J.Myrberg）描述了类似的连续分叉。洛斯阿拉莫斯国家实验室的梅特罗波利斯、保罗·斯坦和米隆·斯坦证明，倍周期之路并不是只有逻辑斯蒂映射才有，事实上任何抛物线形状的映射都有类似现象。这里“抛物线形状”意指映射的图形有一个隆起——用数学术语说就是“单峰（unimodal）”。

第二条普适性质：费根鲍姆常数

到20世纪70年代，物理学家费根鲍姆（图2.16）的发现让倍周期之路得以在数学界闻名。费根鲍姆用一台可编程的台式计算器算出了倍周

期分叉点的R值表（其中“≈”表示“约等于”）：

$$R_1 \approx 3.0$$

$$R_2 \approx 3.44949$$

$$R_3 \approx 3.54409$$

$$R_4 \approx 3.564407$$

$$R_5 \approx 3.568759$$

$$R_6 \approx 3.569692$$

$$R_7 \approx 3.569891$$

$$R_8 \approx 3.569934$$

$$\cdots$$

$$R_\infty \approx 3.569946$$

这里R_1 对应周期2^1 （=2），R 2对应周期2^2 （=4），R_n 对应周期2^n 。符号∞（“无穷大”）用来标志混沌的出现——周期为无穷大的轨

道。

费根鲍姆注意到，随着周期增大，R值之间的距离越来越近。这意味着随着R的增大，分叉之间的间隔越来越短。在图2.15的分叉图中可以看到这一点。费根鲍姆用这些R值计算了分叉靠近的速度，也就是R值的收敛速度。他发现速度约等于常数4.6692016。这意味着随着R值增加，新的周期倍增比前面的周期倍增出现的速度快大约4.6692016倍。



图2.16　费根鲍姆（美国物理学会西格尔图像档案，当代物理藏品）

这很有趣，但还不至于让人震惊。当费根鲍姆研究了其他一些映射后——逻辑斯蒂只是研究过的映射之一——事情变得更有趣了。我在前面提到，在费根鲍姆进行这些计算之前的几年，他在洛斯阿拉莫斯的同事梅特罗波利斯、保罗·斯坦和米隆·斯坦就证明了所有单峰映射都会有类似的倍周期现象。费根鲍姆下一步做的就是计算其他单峰映射的收敛速度。他先算了正弦映射，正弦映射与逻辑斯蒂映射相似，不过用的是正弦函数。

费根鲍姆重复了前面的步骤：计算正弦映射的倍周期分叉点的R值，然后计算这些值的收敛速度。他发现收敛速度为4.6692016。

费根鲍姆很吃惊，速度是一样的。他又检验了其他单峰映射，结果还是一样。所有人，包括费根鲍姆自己，都根本没有想到会是这样。发现这个结果后，费根鲍姆接着又从理论上解释了为何常数4.6692016具有普适性——对所有单峰映射都成立。这个数现在被称为费根鲍姆常数。常数的理论解释使用了一种复杂的数学技巧——重正化（renormalization）。重正化最初是从量子力学中发展出来，后来又被应用到另一个物理学领域：相变和其他“临界现象”的研究。费根鲍姆将其引入了动力系统理论，[31] 并成为理解混沌的基石。

后来发现这并不仅仅是数学现象。费根鲍姆做出这个发现之后，他的理论在多个物理动力系统的实验中得到了证实，包括流体、电路、激光和化学反应。在这些系统中都发现了倍周期分叉，也用类似的方法计算了费根鲍姆常数。在这些实验中很难准确测量分叉点的R值，但即使这样，实验得到的费根鲍姆常数也仍然在接近4.6692016的误差范围之内。这很让人印象深刻，因为费根鲍姆的理论在算出这个数时只涉及数学，没涉及物理。正如费根鲍姆的同事卡达诺夫（Leo Kadanoff）所说的，这是“一个科学家所能遇到的最好的事情，[32] 头脑中想到的东西在自然界中得到了完美的印证”。

气象这样的大尺度的系统很难直接做实验，因此没有人在大尺度系统中直接观察到倍周期分叉或混沌。不过，一些气象计算机模型 [33] 却表现出了通往混沌的倍周期之路，另外电力系统、心脏、行星等系统的计算机模型中也有类似发现。

关于这个故事还有一件让人吃惊的事情。同许多重要的数学发现一样，几乎在费根鲍姆做出他的发现同时，另一个研究小组也独立发现了这个规律。这个小组是法国科学家科雷特（Pierre Coullet）和特雷瑟（Charles Tresser），他们也用重正化技术研究了倍周期分叉，[34] 并且发现了单峰映射的普适常数4.6692016。费根鲍姆也许的确是第一个发现者，并且向科学界广泛而清晰地传播了这个结果，所以这个成就大部分被归功于他。不过在许多科技文献中，也称这个理论为“费根鲍姆—科雷特—特雷瑟理论”，称费根鲍姆常数为“费根鲍姆—科雷特—特雷瑟常数”。在书中还有几个这样的例子，都是在思想条件成熟时同时独立做出发现。

混沌思想带来的革命

在这一章我们看到，混沌的发现使得科学的许多核心原则被重新加以思考。这里我总结一下这些新思想，19世纪的科学家几乎没人会相信这些。

◆看似混沌的行为有可能来自确定性系统，无须外部的随机源。

◆一些简单的确定性系统的长期变化，由于对初始条件的敏感依赖性，即使在原则上也无法预测。

◆虽然混沌系统的具体变化无法预测，在大量混沌系统的普适共性中却有一些“混沌中的秩序”，例如通往混沌的倍周期之路，以及费根鲍姆常数。因此虽然在细节上“预测变得不可能”，但在更高的层面上混沌系统却是可以预测的。

总的来说，变化、难以预测的宏观行为是复杂系统的标志。动力系统理论为刻画其行为提供了数学词汇表，例如分叉、吸引子以及系统变化方式的普适特性。这些词汇在复杂系统的研究中频繁出现。

逻辑斯蒂映射是种群数量增长的简化模型，但是对其以及类似模型的详细研究却带来了对秩序、随机和可预测性的重新认识。这证明了理想模型（idea models）的力量——这些模型很简单，用数学或计算机就足以进行研究，但是又抓住了自然界复杂系统的本质。理想模型在这本书中，乃至整个复杂系统科学中都扮演了重要角色。

刻画复杂系统的动力学还只是理解它的第一步。我们还要理解这些动力系统如何被用在生命系统中以处理信息和适应环境变化。后三章会针对这些主题给出一些背景知识，然后我们再来看看从动力学中得到的思想如何与信息论、计算和进化结合起来。

第3章　信息

我认为，熵增定律 [35] ——热力学第二定律——在自然界的定律中具有至高无上的地位……如果你的理论被发现违背了热力学第二定律，你就一点希望都没有，结局必然是彻底崩塌。

——爱丁顿爵士（Sir Arthur Eddington），

《物理世界的本性》（Te Nature of the Physical World ）

讨论复杂系统时经常会说到“自组织”：例如，行军蚁搭建的桥；萤火虫的同步闪动；经济系统中相互维系的市场；干细胞发育成特定的器官——这些都是自组织的例子。与通常情形中的有序消退、无序（熵）增长相反，这里是有序从无序中产生。

复杂系统科学最关注的问题就是这种逆熵的自组织系统是如何可能的。不过要着手这个问题，还要先了解一下什么是“有序”和“无序”，以及人们如何看待对这种抽象性质的度量。

许多复杂系统学家用信息的概念来刻画和度量有序和无序、复杂性和简单性。免疫学家科恩（Irun Cohen）曾说，“复杂系统比简单系统更能接收、存储和利用信息”。 [36] 经济学家贝哈克（Eric Beinhocker）写道，“进化不仅只会用DNA要把戏， [37] 对所有能处理和存储信息的系统也可以”。物理学家盖尔曼（Murray Gell-Mann）在讨论复杂系统理论时则说，“虽然它们的物理属性很不相同， [38] 但它们处理信息的方式却是类似的。这个共性也许是对它们进行研究最好的起点”。

但是“信息”到底是什么呢？

信息是什么

现在“信息”一词随处可见：信息革命、信息时代、信息技术（常常简化为IT）、信息高速公路，诸如此类。信息在口语中被用来泛指所有表示知识或事实的媒介：报纸、书籍，我母亲在电话里唠叨家里的亲人，还有现在大行其道的万维网。专业点说，信息描述了一大类现象，从在万维网上通过光纤传送的信号，到大脑中在神经元之间传递的微小分子。

在第1章中提到的那些复杂系统的例子无一例外都涉及以各种形式交流和处理信息。进入计算机时代后，科学家们开始想到信息的传递和计算不仅仅发生在电子电路中，在生命系统中也同样存在。

要理解这些系统中的信息和计算，首先当然要对信息和计算这两个术语的意义有精确的定义。两者都是到20世纪才在数学上被定义。让人吃惊的是，两者居然都是从19世纪末的一个物理难题发展而来，这个难题中有个非常聪明的“小妖”，它似乎不用耗费任何能量就能做很多事情。这个难题曾让物理学家们非常担心，以为他们的基本定律可能哪里错了。信息的概念是如何拯救这一切的呢？在了解这些之前，我们先要了解一点关于能量、功和熵等物理概念的背景。

能量、功、熵

对于信息的科学研究始自热力学，热力学描述能量以及其与物质的相互作用。19世纪的物理学家认为宇宙是由物质（固体、液体、气体，等等）和能量（热能、光能、声能，等等）组成。

能量大致上可以定义为系统“做功”的潜力，这符合我们对能量的直观感觉，特别是在这个精力十足的工作狂的时代。英语中能量（energy）一词源自古希腊语中的energia，字面意思是“工作”。不过在物理学中，对一个物体做的“工作”有特定的含义：对物体施加力的大小乘以物体沿力的方向前进的距离。

打个比方，假设你的车在路上抛锚了，你不得不自己把车推到最近的加油站。用物理学的话讲，你做的功等于你推车的力的大小乘以到加油站的距离。在推车的过程中，你将你体内储存的能量转化成了车的动能，而转化的能量就等于所做的功加上轮子与地面摩擦消耗的热量以及你自己体温升高所耗费的热量。这个热量损失可以用熵度量。熵是对不

能转化成功的能量的度量。“熵（entropy）”一词源自另一个古希腊词汇——“trope”——意思是“变成”或“转化”。

在19世纪末，两条关于能量的基本定律——也就是热力学定律——被发现了。这些定律所针对的是“封闭系统”——它们与外界没有能量交换。

第一定律：能量守恒。宇宙中的总能量守恒。能量可以从一种形式转化成另一种形式，比如从体内储存的能量转化成推车的动能加上消耗的热能。但是能量既不能被创生也不能被消灭。因此说是“守恒的”。

第二定律：熵总是不断增加直至最大。系统总的熵会不断增加，直至可能的最大值；除非通过外部做功，否则它自身永远也不会减少。

你可能曾注意过，房间不会自己变干净，饮料如果泼到地上，永远也不会回到杯子里。要想将无序变成有序，就得额外做功。

此外，能量转化的时候，比如前面推车的例子，总是会产生一些不能做功的热能。这也就是为什么没法将你家冰箱后面产生的热量转化成电力再来驱动你的冰箱。这也解释了为何永动机是不可能的。

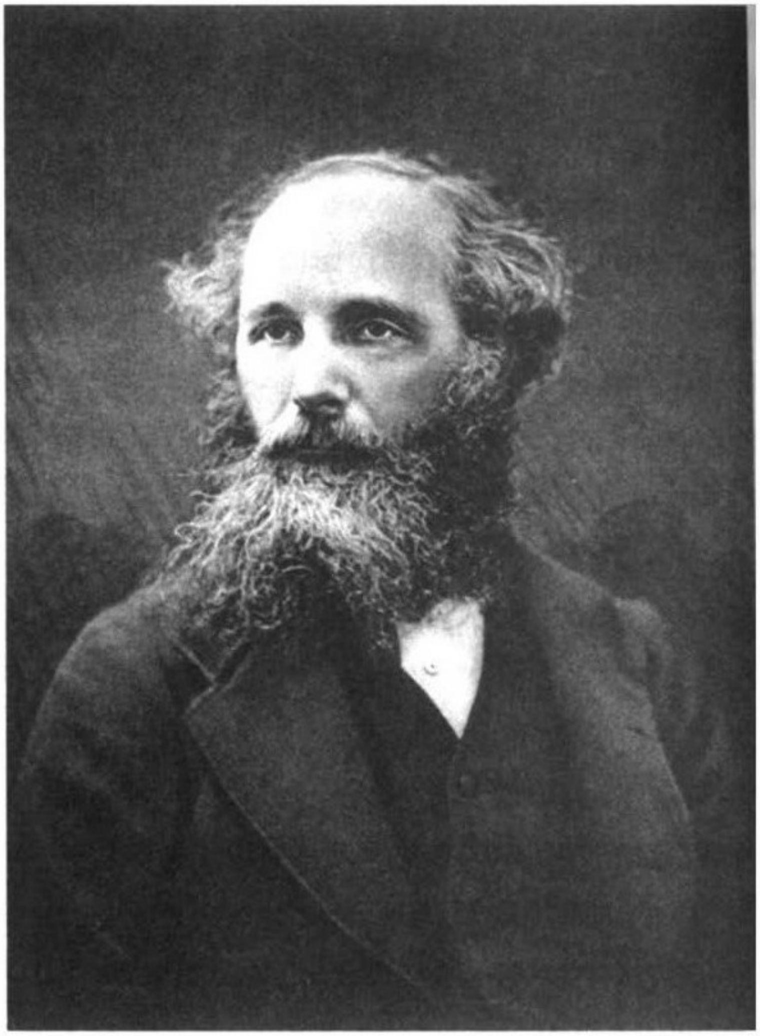
热力学第二定律被认为是定义了“时间之箭”，因为它证明了存在时间上不可逆的过程（比如，热量自发地回到你的冰箱，并转化成电能进行制冷）。“未来”可以定义为熵增的时间方向。有趣的是，热力学第二定律是唯一区分过去和未来的基本物理定律。其他物理定律在时间上都是可逆的。比如，假设可以将电子等基本粒子的相互作用拍成电影，然后给物理学家播放这段电影。如果将电影倒放，然后问物理学家哪个版本是“真实”版本。物理学家肯定猜不出来，因为不管是正放还是倒放，其中的相互作用都没有违反物理定律。这就是可逆的含义。但是如果你用红外胶片拍下冰箱释放热量的过程，然后正放和倒放，物理学家将能辨别出正放的那个是“正确的”，因为遵守了第二定律，而倒放的则没有遵守。这也就是不可逆的含义。为什么第二定律会与众不同呢？这个问题很深奥。就像物理学家罗斯曼（Tony Rothman）所指出的，“为什么第二定律能区分过去和现在，[39] 而其他自然定律却不能？这也许是物

理学中最大的谜团”。

麦克斯韦妖

英国物理学家麦克斯韦（James Clerk Maxwell）提出了著名的麦克斯韦方程，从而统一了电学和磁学。他是当时世界上最受尊敬的科学家，也是古往今来最伟大的科学家之一。

1871年，麦克斯韦在《论热能》（*Theory of Heat* ）一书中提出了一个难题，题为“热力学第二定律的局限”。麦克斯韦假设有一个箱子被一块板子隔成两部分，板子上有一个活门，如图3.1所示。活门有一个“小妖”把守，小妖能测量气体分子的速度。对于右边来的分子，如果速度快，他就打开门让其通过，速度慢就关上门不让通过。对于左边来的分子，则速度慢的就让其通过，速度快的就不让通过。一段时间以后，箱子左边分子的速度就会很快，右边则会很慢，这样熵就增加了。



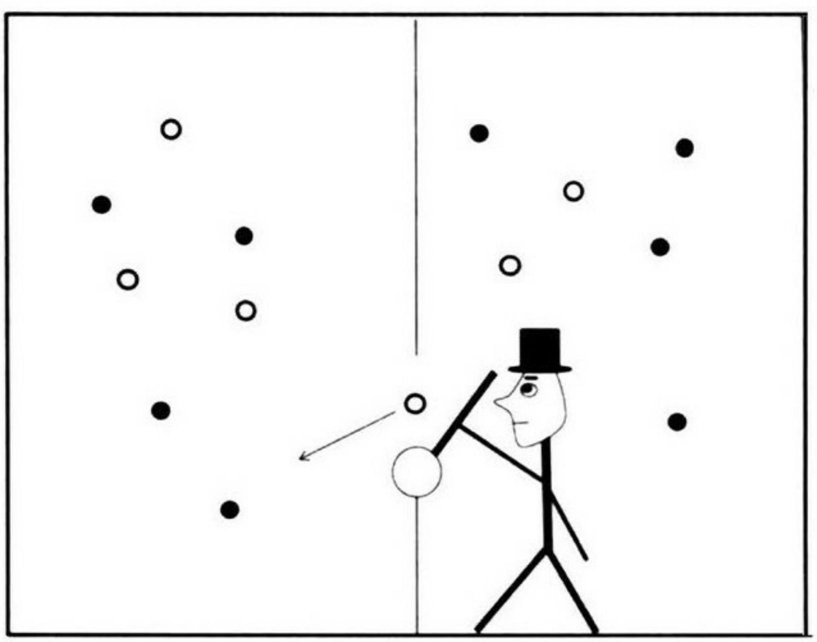


图3.1　上图：麦克斯韦（1831—1879）（美国物理学会西格尔图像档案）。下图：麦克斯韦妖会在快分子（白色）通往左边时和慢分子（黑色）通往右边时打开门

根据热力学第二定律，要减少熵就得做功。小妖又做了什么功呢？当然，他开门关门无数次。但是麦克斯韦假设了小妖使用的门既无质量也无摩擦，因此开门关门要不了多少功，可以忽略不计（对这种门提出了可行的设计）。那么小妖还做了其他的功吗？

麦克斯韦的回答是没有：“热系统（左边）变得更热，[40] 冷系统（右边）变得更冷，然而却没有做功，只有一个眼光锐利、手脚麻利的智能生物在工作。”

为什么没做功，熵也减少了呢？这岂不是违反了热力学第二定律？麦克斯韦的小妖难住了19世纪末和20世纪初许多杰出的头脑。麦克斯韦自己的回答是第二定律（熵随时间增加）根本就不是一条定律，而是在大量分子情形下成立的统计效应，在个体分子尺度上并不必然成立。

但是当时和后来许多物理学家都强烈反对。他们认为第二定律绝对没错，肯定是那个小妖玩了猫腻。既然熵减少了，肯定以某种难以确定的方式做了功，否则不可能。

很多人都想解决这个悖论，但是直到60年后这个问题才被圆满解决。1929年，突破出现了：杰出的匈牙利物理学家西拉德（Leo Szilard）提出，做功的是小妖的“智能”，更精确地说，是通过测量获取信息的行为。

西拉德（图3.2）是第一个将熵与信息联系起来的人，这个关联后来成了信息论的基础和复杂系统的关键思想。西拉德写了一篇题为“热力学系统在智能生物的干预下的熵的减少”的著名论文[41]，文中西拉德认为测量过程（小妖要通过测量获取“比特”信息，比如趋近的分子速度是慢是快）需要能量，因此必然会产生一定的熵，数量不少于分子变得有序而减少的熵。这样由箱子、分子和小妖组成的整个系统就仍然遵守热力学第二定律。



图3.2　西拉德（1898—1964）（美国物理学会西格尔图像档案）

西拉德在此过程中也顺便定义了信息比特的概念——通过回答是/否（对小妖是“快/慢”）获得的信息。他可能是第一个这样做的人。

现在回过头来看，获取信息需要额外做功可能是很显然的事情，起码不那么让人吃惊。但是在麦克斯韦的时代，甚至到60年后西拉德写文章的时候，人们仍然强烈倾向于将物理和精神过程视为完全独立。也许正是这种牢固的直觉使得像麦克斯韦这样睿智的人也没有看出小妖的“智能”或“观测能力”对箱子—分子—小妖系统的热力学有影响。直到20世纪发现“观察者”在量子力学中扮演了关键角色之后，信息与物理的关系才开始被理解。

西拉德的理论后来由法国物理学家布里渊（Leon Brillouin）和伽柏（Denis Gabor）进行了扩展和一般化。此后许多科学家都认为，布里渊的理论彻底揭示了测量是如何产生熵，从而终结了小妖。

然而，事情还没有结束。在西拉德的论文发表50年后，西拉德和布里渊的论证都被发现有一些漏洞。20世纪80年代，数学家班尼特（Charles Bennett）证明，[42] 有非常巧妙的方式可以观察和记住信息——对小妖来说，也就是弄清分子是快是慢——而不用增加熵。班尼特的证明成了可逆计算（reversible computing）的基础，他证明在理论上可以进行任何计算而不用耗费能量。班尼特的发现似乎意味着小妖又回来了，因为测量可以不用耗费能量。不过，班尼特认为，物理学家兰道（Rolf Landauer）在20世纪60年代做出的一项发现可以挽救热力学第二定律：并不是测量行为，而是擦除记忆的行为，必然会增加熵。擦除记忆是不可逆的；如果被擦除了，那么一旦信息没有了，不进行额外的测量就无法恢复。班尼特证明，小妖如果要工作，到一定的时候就必须擦除记忆，如果这样，擦除的动作就会产生热，增加的熵刚好抵消小妖对分子进行分选而减少的熵。

兰道和班尼特弥补了西拉德论证的漏洞，但思路仍然是一致的：小妖测量和进行判断时（必然会进行擦除），不可避免地会增加熵，从而热力学第二定律仍然成立。（不过仍然有一些物理学家不认可兰道和班尼特的论证，小妖的问题依然存在争议。[43] ）

麦克斯韦发明小妖是将其作为一个简单的思维实验，以证明热力学

第二定律不是一条定律，而只是统计效应。然而，同其他许多优秀的思维实验一样，小妖的影响很深远；对小妖难题的解决成为两个新领域的基础——信息论和信息物理学。

统计力学提要

在前面我将“熵”定义为对无法做功而只能转换成热的能量的测量。这个熵的概念最初是由克劳休斯（Rudolph Clausius）于1865年定义的。在克劳休斯的年代，热被认为是某种可以从一个系统流向另一个系统的流质，而温度则是系统受热流影响的一种属性。

此后数十年里，科学界开始出现一种新的关于热的观念：系统是由分子组成，而热则是分子运动——或者说动能——的产物。这种新观念主要归功于玻尔兹曼（Ludwig Boltzmann，图3.3），他创建了一门新学科，现在被称为统计力学。



图3.3　玻尔兹曼（1844—1906）（美国物理学会西格尔图像档案，西格尔收藏）

统计力学认为宏观尺度上的属性（例如热）是由微观属性产生（例如无数分子的运动）。比如，想象房间里充满了运动的空气分子。经典力学分析是确定每个分子的位置和速度，以及作用在分子上的力，并根据这些确定每个分子未来的位置和速度。当然，如果有500亿亿个分子，要解出来可得花不少时间——实际上是完全不可能的，并且根据量子力学，在原则上也不可能。而统计力学的方法则不关心各个分子具体的位置、速度以及未来的变化，而是去预测大量分子整体上的平均位置和速度。

简而言之，经典力学试图用牛顿定律分析所有的单个微观对象（例如分子）。而热力学则只给出了宏观现象——热、能量和熵——的定律，没有说明微观分子是这些宏观现象的源头。统计力学则在两个极端之间搭建了一座桥梁，解释了宏观现象是如何从对大量微观对象的整体上的统计产生。

统计方法有一个问题——它只给出系统的可能行为。例如，如果房间里的空气分子随机运动，那么它们将极有可能扩散到整个房间，从而保证我们所有人都可以呼吸到空气。我们预计会这样，并且生命维系于此，而且也从没有失败。然而，根据统计力学，由于分子是随机运动，这样就存在一个极小的概率在某个时间分子都飞到一个角落里。然后那个角落里的人会被高气压压死，而我们其他人则会窒息而死。不过据我所知，这样的事情还从未发生过。这并不违反牛顿定律，只是极为不可能。玻尔兹曼认为，如果有足够多的微观对象进行平均，他的统计方法就几乎一直都能给出正确答案，而事实上也确实如此。但是在玻尔兹曼的时代，大部分物理学家都只接受[44]绝对正确的物理定律，“几乎一直”正确的物理定律是不会被接受的。此外，玻尔兹曼认为存在分子和原子这样的微观对象也让他的同行们感到不可理喻。玻尔兹曼于1906年自杀离世，有人认为这是大多数科学家对他的思想排斥所导致的。他死后不久，他的思想就被广泛认同了；现在他被认为是历史上最伟大的科学家之一。

微观态与宏观态

在充满空气的房间中，在任意时刻每个分子都有特定的位置和速

度，只是无法具体测量。在统计力学的术语中，特定分子集合在某一时刻的位置和速度称为那个时刻的微观状态。对于充满了随机飞舞的分子的房间，最可能的微观状态类型就是空气分子均匀地充满整个房间。而最不可能的微观状态就是空气分子紧紧地聚到一个地方。这看上去显而易见，但是玻尔兹曼注意到这是因为分子均匀分布的微观状态比聚到一起的微观状态要多得多。

这种情形有点类似吃角子老虎（图3.4）。假设三幅图片可能为“苹果”“橙子”“樱桃”“梨”或“柠檬”。你投个硬币进去，让老虎机转起来。图片存在不同（你输钱）的可能性比图片全部相同（你大赢一笔）的可能性要大得多。现在假设老虎机有500亿亿种图片，要让所有图片都相同就类似于让所有分子都聚到一点的情形，可能性基本为零。

系统的宏观状态就是微观状态的类型，例如，“所有图片都相同——你赢”相对“图片不完全相同——你输”，或者“分子聚集到一起——我们窒息”相对“分子均匀分布——我们能呼吸”，一个宏观状态能对应许多不同的微观状态。玩老虎机时，有各种由不同图片组成的微观状态，这些微观状态都对应于同一个宏观状态“你输”，而只有不多的微观状态对应宏观状态“你赢”。这就是为什么赌场能挣大钱的原因。温度也是宏观状态——它与许多不同的微观状态相对应，各微观状态的分子平均速度恰好对应相同的温度。

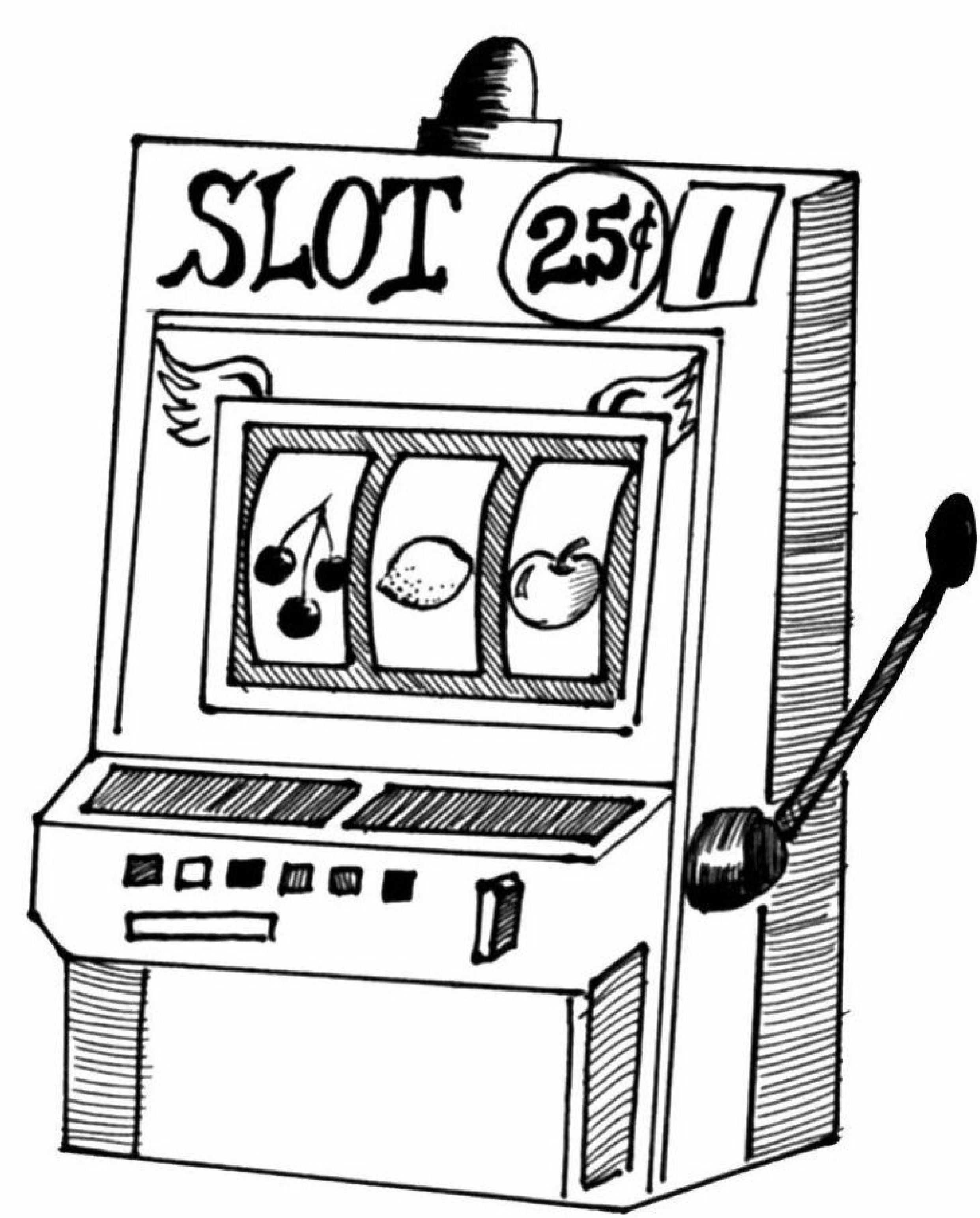


图3.4　有三个旋转图片的老虎机，说明微观状态和宏观状态的概念（David Moser绘制）

根据这些思想，玻尔兹曼将热力学第二定律解释为封闭系统更有可

能处于可能性大的宏观状态。这听起来像是废话，不过在当时这种想法却相当离经叛道，因为涉及了概率的概念。玻尔兹曼将宏观状态的熵定义为[45]其对应的微观状态的数量。例如，图3.4的老虎机中，图片可以是“苹果”“橙子”“樱桃”“梨”或“柠檬”，这样就总共有125种可能的组合（微观状态），其中有5种对应于“所有图片都相同——你赢”的宏观状态，120种对应于“图片不完全相同——你输”的宏观状态。后一种宏观状态的玻尔兹曼熵明显高于前一种。

玻尔兹曼熵遵守热力学第二定律。除非做功，否则玻尔兹曼熵会一直增加，直到到达最大可能熵的宏观状态。玻尔兹曼证明，在许多情形下，他对熵的简单定义与克劳休斯的定义等价。

玻尔兹曼熵的公式[46]被刻在维也纳玻尔兹曼的墓碑上（图3.5），现在这个方程已经成为物理学的基石。

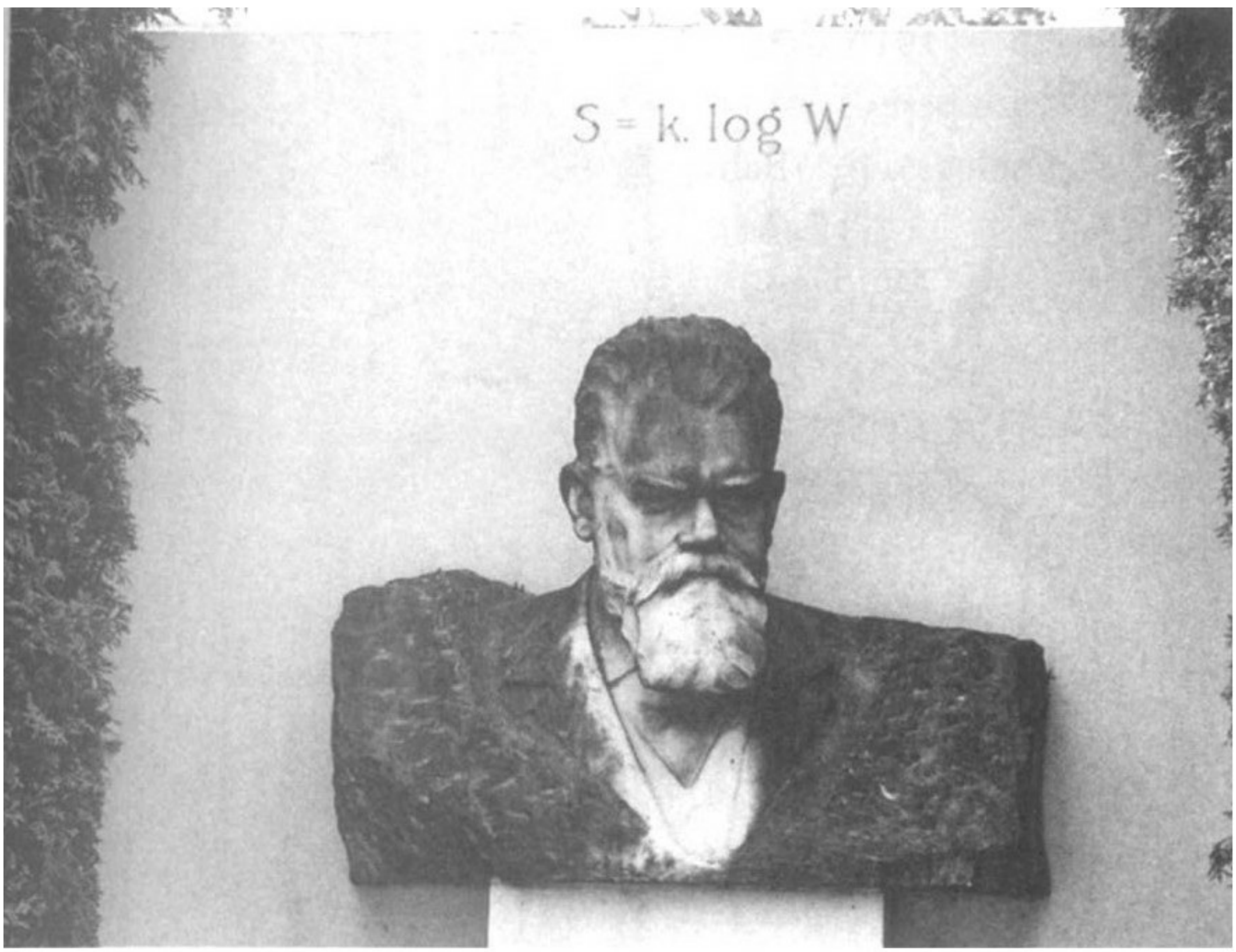


图3.5　玻尔兹曼的墓碑，维也纳（Martin Roell提供图片）

香农信息

科学上许多最基本的思想都是由技术进步促进。19世纪的热力学研究就是由改进蒸汽机时遇到的挑战驱使。而数学家香农（Claude Shannon，图3.6）发展信息论也是受20世纪的通信革命推动，尤其是电报和电话的发展。1940年，香农改进了玻尔兹曼的思想，以适用于更为抽象的通信领域。香农在美国电话电报公司（AT&T）贝尔实验室工作。AT&T当时面临的最重要的问题就是如何通过电报和电话线快速有效地传送信息。

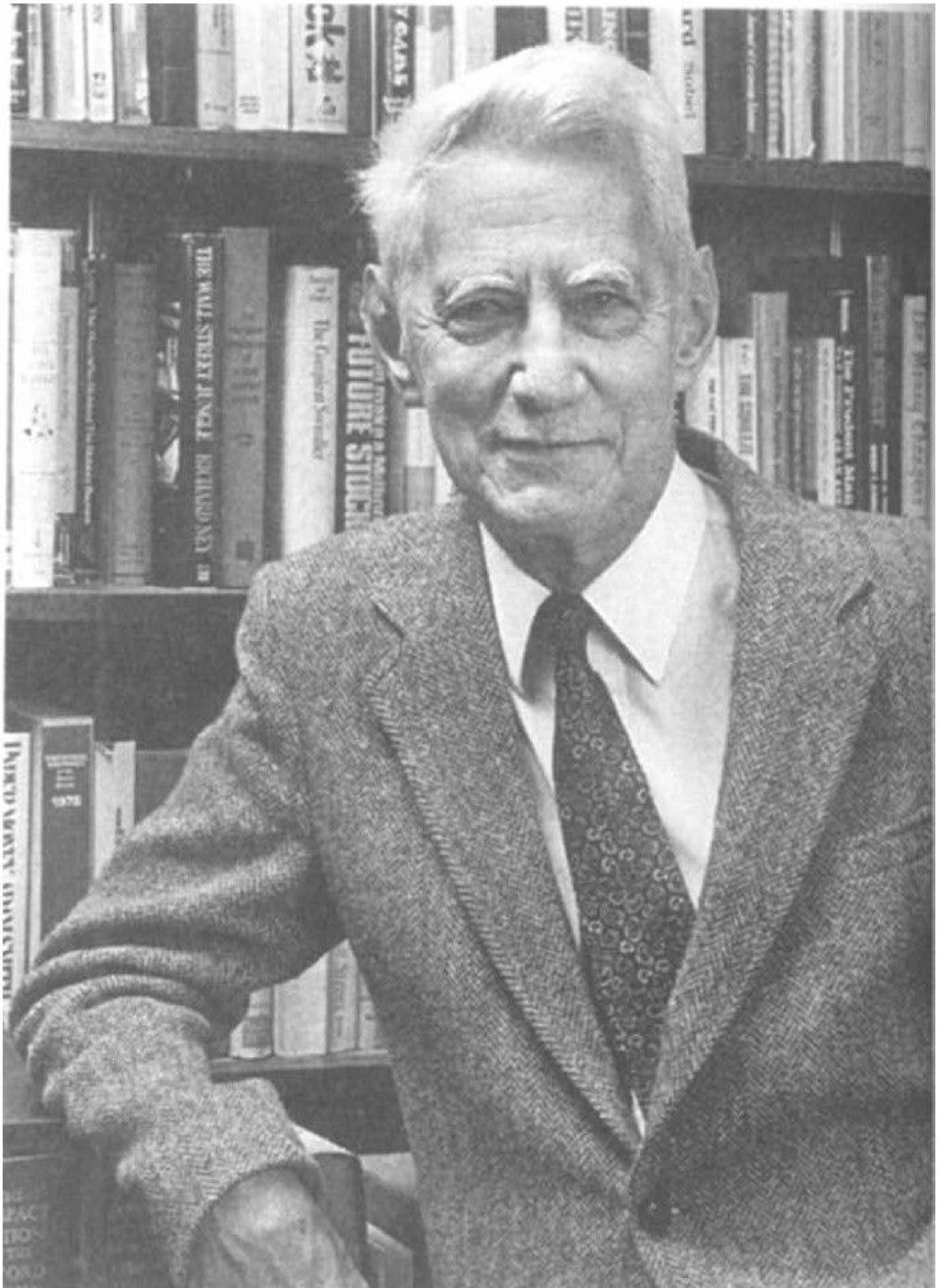
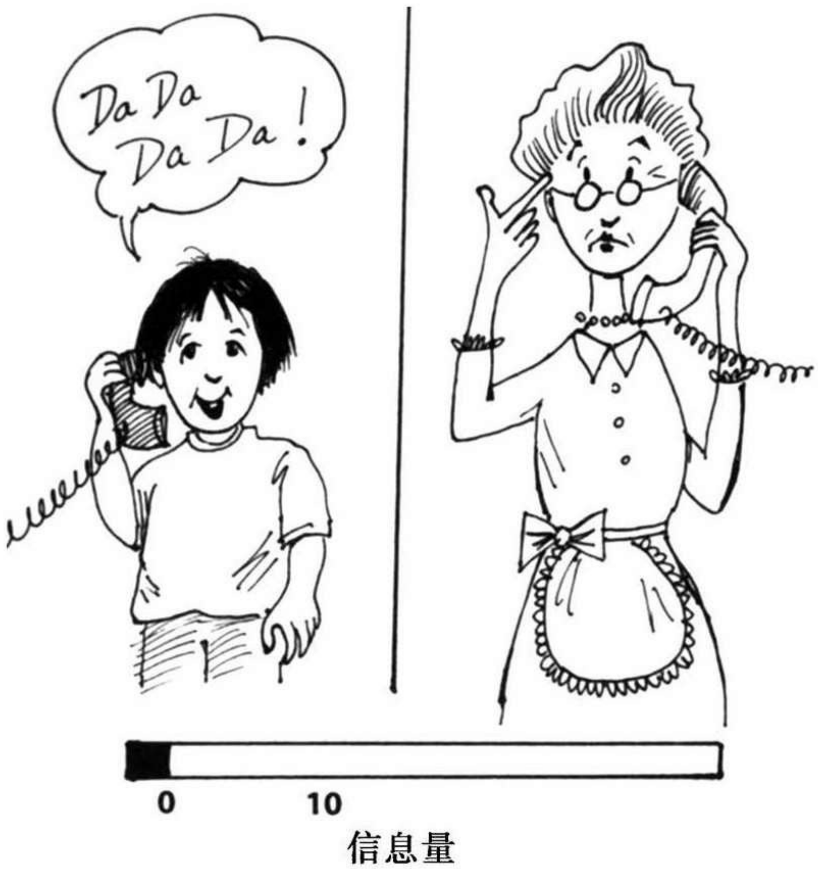


图3.6　香农（1916—2001）（经朗讯公司贝尔实验室许可使用）

香农从数学上解决了这个问题，从而开创了一个新领域——信息论。1948年，香农发表了论文“通信的数学理论”，[47] 在文中香农给出了信息的一个狭义定义，并且证明了一个非常重要的定理，定理给出了通过给定通道传输的最大可能传输率，无论信道是否存在噪声。这个最大传输率就是信道容量（channel capacity）。

香农的信息定义中有一个发送者向接收者发送信息。例如图3.7有两个发送者通过电话与接收者交谈的例子。发送者说的每个词都是香农意义上的信息。电话并不理解所说的词，而只是传送编码声音的电脉冲，香农对信息的定义也完全忽略信息的意义，而只考虑发送者向接收者发送信息的速度。

香农问：“发送者传送了多少信息给接收者呢？”与玻尔兹曼的思想类似，香农将宏观状态（这里是发送者）的信息定义为可以由发送者发送的可能微观状态（可能信息的集合）的数量的函数。我的儿子尼可还在蹒跚学步时，我会让他通过电话同奶奶讲话。他喜欢讲电话，不过只会说一个词——“Da”。他发给奶奶的信息是“Da Da Da Da Da……”换句话说，尼可的宏观状态只有一种可能的微观状态（“Da”序列），因此虽然这个宏观状态很有趣，但信息量却为零。奶奶知道听到的会是什么。我的儿子杰克两岁了，他也喜欢讲电话，不过他的词汇量大些，因此会告诉奶奶他干的事情，经常让奶奶对他讲的话吃惊。显然发送者杰克的信息量要多得多，因为可能的微观状态——即各种不同的信息组成的集合——要多得多。


Da Da
Da Da !
0
10
信息量


嗨，奶奶！
我在扮超人！
0
10
信息量

图3.7　上图：尼可同奶奶交谈的信息量（为零）。下图：杰克同奶奶的交谈有更多的信息量（David Moser绘制）

香农对信息量的定义与玻尔兹曼对熵更一般化的定义几乎一样。在1948年的经典文章中，香农用信息源的熵定义信息量（这个熵的概念通常被称为香农熵，以区别于玻尔兹曼给出的熵的定义）。

人们有时候将香农的信息量定义描述为接收者在接收信息时体验到的“平均惊奇度”，其中“惊奇”意指接收者对于发送源将要传送的信息的“不确定度”。奶奶对杰克所说的肯定会比对尼可所说的更觉得惊奇，因为她完全知道尼可会说什么，却不那么容易知道杰克会说什么。因此杰克所说的给她的平均“信息量”要比尼可说的多。

总体上，根据香农的理论，信息可以是通信的任何单位，可以是一个字母、一个词、一句话，甚至是一个比特（0或1）。发送源的熵（信息量）用信息的可能性定义，而与信息的“意义”无关。

香农的结果在许多领域都有应用。最广为人知的应用就是编码理论，研究数据压缩问题和可靠传输的编码方法。编码理论对电子通信的所有领域几乎都有影响：移动电话、计算机网络、全球定位系统，等等。

信息论也是密码学和新兴的生物信息学的基础，生物信息学通过分析基因序列的模式测量熵等信息论度量。信息论也被应用到语言和音乐的分析，以及心理学、统计推断和人工智能等领域。虽然信息论受到热力学和统计力学熵的概念启发，信息论对物理学的各领域是否有反向影响还有争议。1961年，通信工程师和作家皮尔斯（John Pierce）开玩笑说：“让通信理论和物理学联姻的努力[48]有趣却没什么结果。”一些物理学家认同他的观点。不过，一些基于香农信息论的物理学新思路（例如量子信息论和信息物理学）正不断发展。

在后面你会看到，熵、信息量、交互信息、信息动力学等信息论中的思想在对复杂性概念的定义和对各种类型复杂系统的刻画中扮演了重要而富有争议的角色。

第4章　计算

Quo facto, [49] quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo dicere: Calculemus! （如果产生了争议，哲学家们用不着像会计师一样相互争执，他们只需要掏出纸和笔，然后说：来，演算一下。）

——莱布尼茨（转译自罗素译文）

在普通人眼里，计算就是计算机做的事情，电子表格、文档处理、电子邮件，诸如此类。计算机在人们脑海里就是台式电脑或笔记本，里面有电子电路，一般都带有显示器和鼠标，以前还流行用真空管。对于我们自己的大脑，我们也模糊地觉得有点像计算机，有逻辑演算、记忆存储和输入输出。

不过如果你读复杂系统方面的学术文献，你会发现计算这个词的用法蛮奇怪："细胞和组织中的计算" [50] ；"免疫系统的计算" [51] ；"市场的分布式计算的本质和局限" [52] ；"植物中的涌现计算"。 [53] 这样的例子数不胜数。

自从计算机诞生以来，计算的概念已经走过了很长一段时间，现在许多科学家都将计算视为自然界中很普遍的现象。细胞、组织、植物、免疫系统和金融市场显然和计算机的运作方式不一样，那么他们说的计算到底是什么呢？他们又为什么要这样说呢？

在第12章我们会讨论这些问题，在此之前我们先了解一下计算思想的历史以及科学家用来理解自然界复杂系统的计算概念的基础。

什么是计算？什么可以计算

香农的信息定义关注的是消息源的可预测性。不过在现实世界中，信息是用来分析并产生意义的东西，信息被存储，并和其他信息结合，产生结果或行为。总之，信息是用来计算的。

历史上计算的意义变化很大。直到20世纪40年代末，计算都是指手工进行数学运算（小学生称之为"做算术"）。计算员（Computer）就是

做数学运算的人。我以前的老师伯克斯（Art Burks）常和我们说他娶的是“计算机”——指的是第二次世界大战时被征召入伍手工计算弹道的妇女，伯克斯的夫人在遇到他时正是这样一位计算员。

现在计算指的是各式各样的计算机干的事情，另外自然界的复杂系统似乎也干这个。但是计算到底是什么呢？它又能做些什么呢？计算机什么都能算吗？是不是存在原则上的局限性？这些问题都是在20世纪中叶才得到解决。

希尔伯特问题和哥德尔定理

对计算的基础及其局限的研究，导致了电子计算机的发明，但其最初的根源却是为了解决一组抽象（而且深奥）的数学问题。这些问题是德国数学大师希尔伯特（David Hilbert，图4.1）于1900年在巴黎的国际数学家大会上提出来的。

希尔伯特在演讲中提出了世纪之交面临的23个亟待解决的数学问题。其中第2个和第10个问题在后来影响最大。实际上，它们不仅仅是数学内部的问题，它们还是关于数学本身以及数学能证明什么的问题。总的来说，这些问题可以分为三个部分：

1.数学是不是完备的？也就是说，是不是所有数学命题都可以用一组有限的公理证明或证否。

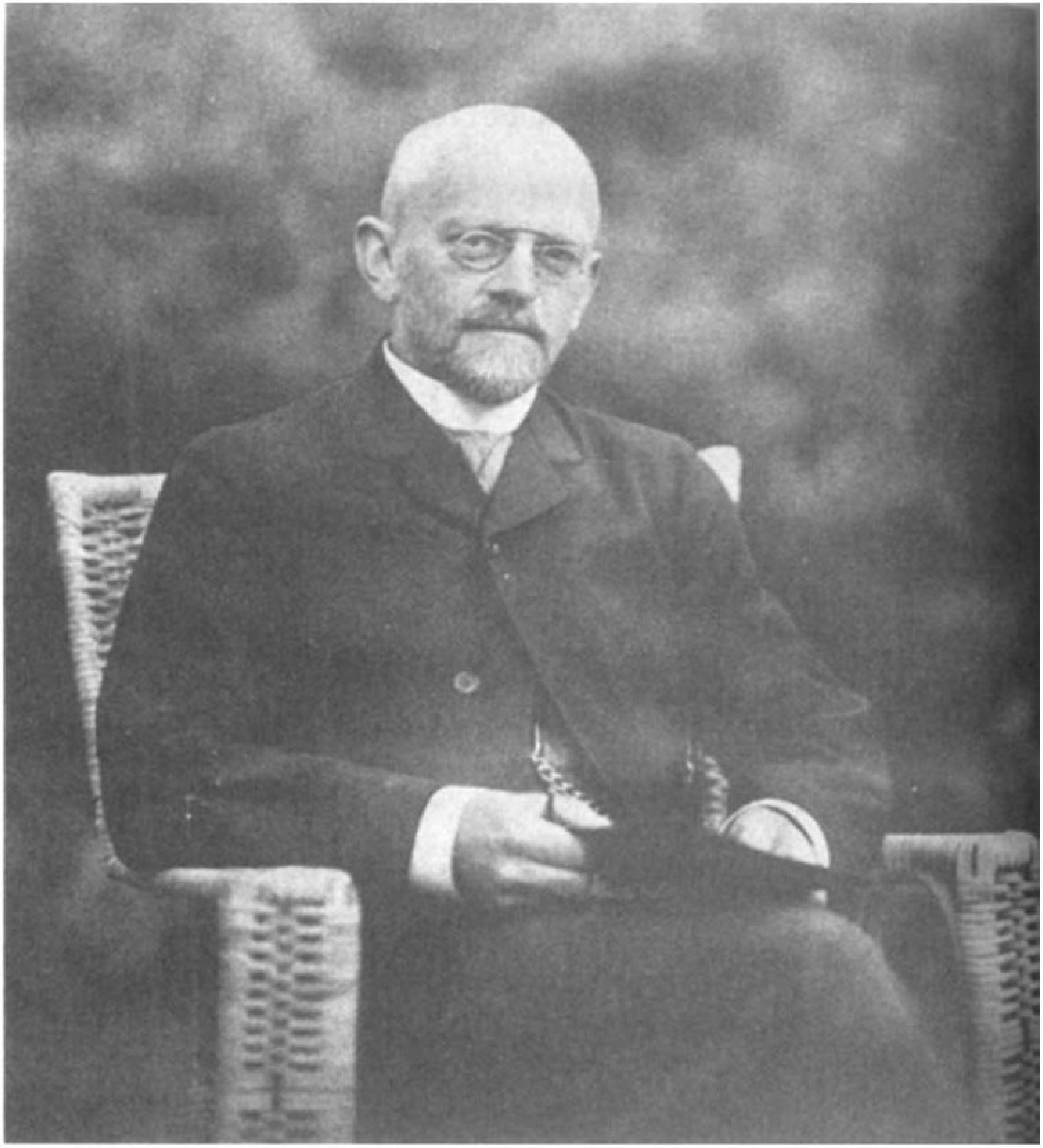


图4.1　希尔伯特（1862—1943）（美国物理学会西格尔图像档案，兰德收藏）

举个例子，还记得中学几何里学过的欧几里得公理吧？记不记得用这些公理可以证明“三角形内角和为180度”这样的定理？希尔伯特的问题是：是不是有某个公理集可以证明所有真命题？

2.数学是不是一致的？换句话说，是不是可以证明的都是真命

题？“真命题”是专业术语，但我在这里用的是直接意义。假如我们证出了假命题，例如1+1=3，数学就是不一致的，这样就会有大麻烦。

3.是不是所有命题都是数学可判定的？也就是说，是不是对所有命题都有明确程序（definite procedure）可以在有限时间内告诉我们命题是真是假？这样你就可以提出一个数学命题，比如“所有比2大的偶数都可以表示为两个素数之和”，然后将它交给计算机，计算机就会用“明确程序”在有限时间内得出命题是“真”还是“假”的结论。

最后这个问题就是所谓的Entscheidungsproblem（“判定问题”），它可以追溯到17世纪的数学家莱布尼茨（Gottfried Leibniz）。莱布尼茨建造了他自己的计算机器，并且认为人类将建造出能判定所有数学命题真假的机器。

这三个问题过了30年都没有解决，不过希尔伯特很有信心，认为答案一定是“是”，并且还断言“不存在不可解的问题”。[54]

然而他的乐观断言并没有维持太久，可以说非常短命。因为就在希尔伯特做出上述断言的同一次会议中，一位25岁的数学家宣布了对不完备性定理的证明，他的发现震惊了整个数学界，这位年轻人名叫哥德尔（Kurt G del，图4.2）。不完备性定理说的是，如果上面的问题2的答案是“是”（即数学是一致的），那么问题1（数学是不是完备的）的答案就必须是“否”。

哥德尔的不完备性定理是从算术着手。他证明，如果算术是一致的，那么在算术中就必然存在无法被证明的真命题——也就是说，算术是不完备的。而如果算术是不一致的，那么就会存在能被证明的假命题，这样整个数学都会崩塌。



图4.2　哥德尔（1906—1978）（照片由普林斯顿大学图书馆提供）

哥德尔的证明很复杂。[55] 不过直观上却很容易解释。哥德尔给出了一个数学命题，翻译成白话就是“这个命题是不可证的”。

仔细思考一下。这个命题很奇怪，它居然谈论的是它自身——事实上，它说的是它不可证。我们姑且称它为“命题A”。现在假设命题A可证，那它就为假（因为它说它不可证），这就意味着证明了假命题——从而算术是不一致的。好了，那我们就假设命题A不可证，这就意味着命题A为真（因为它断言的就是自己不可证），但这样就存在不可证的真命题——算术是不完备的。因此，算术要么不一致，要么不完备。

难以想象这个命题如何转换成用数学语言表述，但是哥德尔做到了——哥德尔证明的复杂和精彩之处就在此，在这里我们不去讨论。

绝大多数数学家和哲学家都坚定地认为希尔伯特问题能被正面解决，这对他们是个沉重的打击。就像数学作家霍吉斯（Andrew Hodges）说的：“这是在研究中惊人的转折，[56] 因为希尔伯特曾以为他的计划将一统天下。对于那些认为数学完美而且无懈可击的人来说，这让人难以接受……”

图灵机和不可计算性

哥德尔干净利落地解决了希尔伯特第一和第二问题，接着第三问题又被英国数学家图灵（Alan Turing，图4.3）干掉了。[57]

1935年，图灵23岁，在剑桥跟随逻辑学家纽曼（Max Newman）攻读研究生。纽曼向图灵介绍了哥德尔刚刚得出的不完备性定理。在理解哥德尔的结果之后，图灵发现了该如何解决希尔伯特的第三问题， 判定问题，同样，他的答案也是“否”[58] 。

图灵是怎么证明的呢？前面说过，判定问题问的是，是否有“明确程序”可以判定任意命题是否可证？“明确程序”指的是什么呢？图灵的第一步就是定义这个概念。沿着莱布尼茨在两个世纪以前的思路，图灵通过构想一种强有力的运算机器来阐述他的定义，这个机器不仅能进行算术运算，也能操作符号，这样就能证明数学命题。通过思考人类如何

计算，他构造了一种假想的机器，这种机器现在被称为图灵机。图灵机后来成了电子计算机的蓝图。



图4.3　图灵（1912—1954）（Photo Researchers公司版权所有©2003，经许可重印）

图灵机概述

如图4.4所示，图灵机由三部分组成：

1.带子，被分成许多方格（或“地址”），符号可以被写入其中或从中读出。带子两头都有无限长。

2.可以移动的读写头，能从带子上读取符号或将符号写到带子上。在任何时候，读写头都处于一组状态中的一个。

3.指示读写头下一步如何做的一组规则。

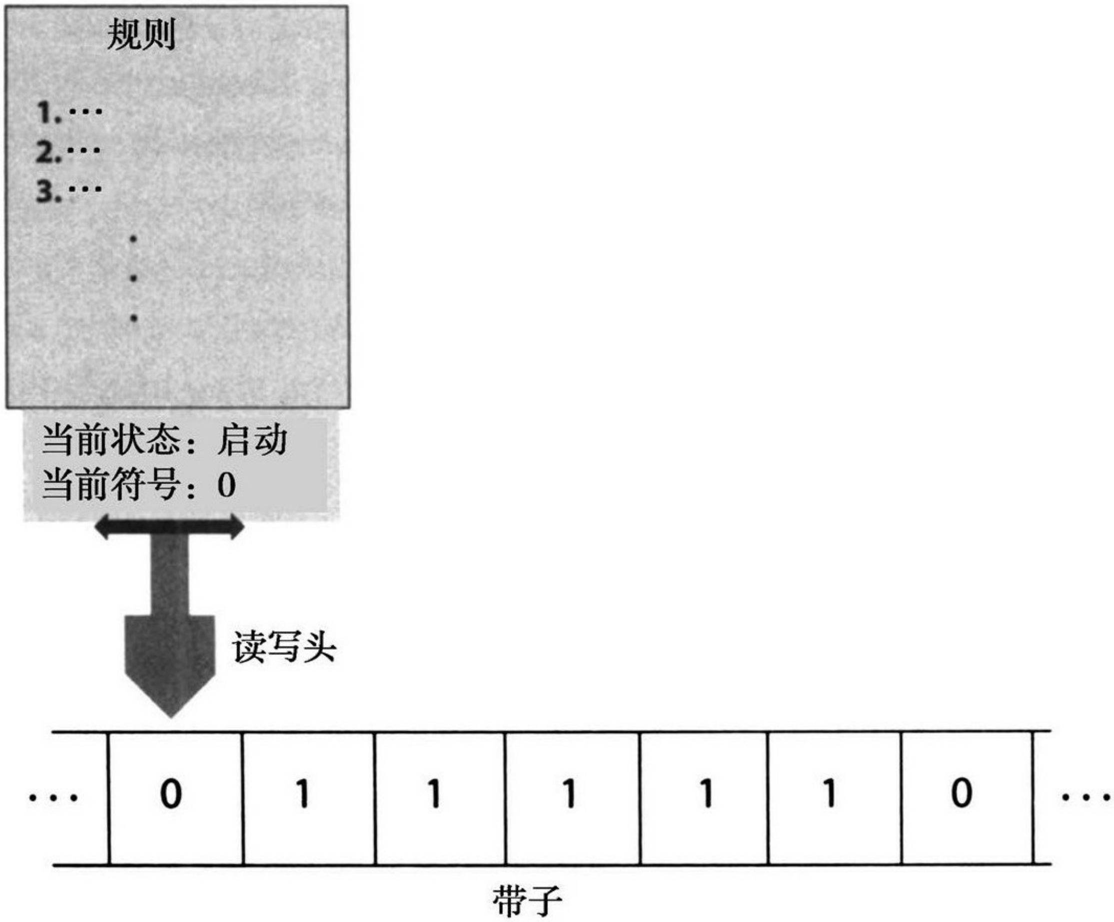


图4.4　图灵机

读写头开始处于特定的开始状态，并停在特定的格子上。每一步，读写头读取当前格子中的符号。然后读写头根据读取的符号和读写头的当前状态按照规则动作。规则决定读写头在当前格子中写入什么符号（替换当前符号）；读写头是向右还是向左移动或是停止不动；以及读写头的新状态是什么。如果读写头进入停机状态，机器就会停下来。

图灵机的输入是机器启动之前写在带子上的符号集合。输出则是停机之后留在带子上的符号集。

一个简单的例子

用一个简单的例子解释一下。为简便起见，假设带子上的符号只有0和1（同真正的计算机一样），而空符号则指格子中为空。我们设计一个图灵机，它读取的带子上有两个0，中间夹着一串1（例如，

011110），然后判断1的个数是奇数还是偶数。如果是偶数，机器最后就在带子上输出一个0（其他格子为空）。如果是奇数，最后就在带子上输出一个1（其他格子为空）。假设带子的输入总是刚好有两个0，中间夹着零个或多个1，其他的格子都为空。

我们的图灵机的读写头得有4种可能状态：启动、偶数、奇数和停机。读写头最初停在最左边的0，处于启动状态。我们编写规则让读写头往右移动，每次一格，用空格替换遇到的0和1。如果读写头在当前格子中读到1，读写头就变成奇数状态，并往右移动一格。如果再次读到1，就变成偶数状态，再往右移动一格。就这样读到1就在偶数和奇数状态之间切换，一直往下。

如果读写头读到0，输入就走完了，这时所处的状态（奇数或偶数）就是想要的结果。然后机器根据状态在当前格子里写1或0，并变成停机状态。

下面是读写头实现这个算法的规则：

1.如果处于启动状态，当前格子为0，就变成偶数状态，把0擦掉，并往右移动一格。

2.如果处于偶数状态，当前格子为1，就变为奇数状态，把1擦掉，并往右移动一格。

3.如果处于奇数状态，当前格子为1，就变成偶数状态，把1擦掉，并往右移动一格。

4.如果处于奇数状态，当前格子为0，就在格子中写入1，并变成停机状态。

5.如果处于偶数状态，当前格子为0，就在格子中写入0，并变成停机状态。

首先在带子上写好输入，然后让图灵机根据规则顺序处理输入，这个过程被称为“根据输入运行图灵机”。

定义为图灵机的明确程序

在上面的例子中，如果输入的格式正确，不管具体的输入是什么（包括一个1也没有的情况，这时视为有偶数个1），根据输入运行图灵机都能确保得出正确的输出。虽然这有点小儿科，但你还是得承认这是一个解决奇偶问题的“明确程序”，每一步都很明确。图灵的第一个目标就是落实明确程序的概念。其中的思想是，对于特定的问题，你可以通过设计一个解决这个问题的图灵机来构造明确程序。这样“明确程序”就定义为图灵机，虽然目前还有些模糊不清。

在对此进行思考时，图灵并没有真的去建造一台机器（虽然后来他这样做了）。他对图灵机的所有思考都是通过纸和笔完成的。

通用图灵机

接下来，图灵又证明了图灵机的一个神奇特性：人们可以设计出一种通用图灵机（称之为U），它可以模拟任何图灵机的运作。U在模拟图灵机M处理输入I时，U处理的带子上不仅包含编码输入I的序列，还包含编码图灵机M的序列。你可能会奇怪图灵机M也能被编码，不过这其实不难。首先，所有规则（同前面那五条差不多）都可以简写成下面这种形式：

—当前状态—当前符号—新状态—新符号—动作—

这样前面的规则1就表示成：

—启动**—0—**偶数**—**空**—**右移**—**

（分隔符“—”不是必需的，只是方便我们读规则。）然后用0/1序列对规则进行编码：例如，启动=000，偶数=001，奇数=010，停机=100。类似的，符号也能编码成0/1序列：例如，符号“0”=000，符号“1”=001，空符号=100。0/1序列可以随意设定，只要与符号一一对应就行。状态和符号——比如启动和“0”——之间使用的0/1序列可以相同；因为我们知道规则的结构形式，据此就能分析出编码的内容。

类似的，也可对动作进行编码，“右移”对应000，“左移”对应111。将分隔符“—”编码成111。这样整条规则都可以编成0/1序列了。编码出来的规则1是这个样子：

111000111000111001111100111000111

其他规则也可以依此编为0/1码。将所有规则的编码连到一起，形成一个长串，就是图灵机M的编码。为了让U模拟M处理I, U最初的带子上既包含I也包含M的编码。U的每一步在带子的输入部分读取I的当前符号，并从带子的M部分读取相应的规则，应用于输入部分；这样就能模拟跟踪M在处理给定输入时的状态和输出。

如果M到达停机状态，U也会停机，并且带子上产生的输出会与M处理给定输入I时产生的输出一样。因此我们说“U在I上运行M”。这里没有讨论U本身的状态和规则，因为比较复杂，但是这种图灵机是肯定可以设计出来的。事实上，现在的可编程计算机正是这样一台通用图灵机：它读取存储的程序，并在存储的输入I上运行这个程序。在图灵证明了存在通用图灵机之后十来年，第一台可编程的计算机就被建造出来了。

图灵对判定问题的解决

再来看看判定问题：是否有明确程序可以判定任意命题是否为真？

通过证明通用图灵机的存在，图灵证明了这个问题的答案是“否”。他意识到图灵机的编码也可以作为另一台图灵机的输入。这就好像用一个计算机程序作为另一个程序的输入。比方说，你可以用WORD写一个程序，存成一个文件，然后用另一个程序（字数统计）处理这个文件，输出你的程序的字数。字数统计程序并不运行你的程序，它只是统计文本文件中的字数。

类似的，你也可以设计出统计输入中1的个数的图灵机M（这个不是很难），然后运行M处理另一个图灵机M′。M会统计M′中的1的个数。当然，在通用图灵机U中，将M的代码置于U的带子的“程序”部分，将M置于带子的“输入”部分，U就能在M上运行M。为了好玩，我们也可以在U的带子的“程序”和“输入”部分都放置M，让M处理它自己的编码！这就好像让字数统计程序来计数它自己的字数。完全没有问题！

带子上的0/1序列既可以作为程序，也可以作为另一个程序的输

入。对熟悉计算机的你来说，这一点可能稀松平常，但是在图灵提出他的证明的年代，这点洞察却是革命性的。

现在我们来看看图灵的证明。

图灵是用反证法证明判定问题的答案为“否”。首先假设答案是“是”，然后证明这个假设会导致矛盾，从而证明答案是相反的。

图灵首先假设，存在明确程序可以判定任意给定命题是否为真。然后他给出了这样一个命题：

图灵命题：图灵机M对于给定输入I会在有限步后停机。

根据前面的假设，将M和I作为输入，有一个明确程序可以判定这个命题是否为真。图灵证明，这个假设会导致矛盾。

我们注意到有一些图灵机永远也不会停机。例如，考虑与前面例子类似的一个图灵机，只有两条规则：

1.如果处于启动状态，读取到0或1，变为偶数状态，并往右移动一格。

2.如果处于偶数状态，读取到0或1，变为启动状态，并往左移动一格。

这个图灵机完全没有错误，但是它永远不会停机。用现在的话说是“死循环”——写程序时经常会碰到的bug。死循环有时候隐藏很深，让你很难发现。

前面假设了有明确程序能够判定图灵命题是否成立，这就等价于说我们能设计一个图灵机来检查死循环。

具体说，这个假设说的是我们能设计一个图灵机H，对于任何给定的图灵机M和输入I，都能在有限时间内判断M对输入I是会停机还是会进入死循环，停不了机。

设计H的问题被称为“停机问题”。注意到H本身必须总是能停机，不管答案是“是”还是“否”。因此H不能通过在I上运行M来做到这一点，

因为有可能M不会停机，从而H也停不了。H必须想其他办法做到这一点。

虽然还不清楚H要如何做，但是我们假设了H是存在的。我们将运行H处理M和I记为H（M，I）。如果M对于I会停机就输出“是”（例如在带子上写一个1），如果M对于I不会停机就输出“否”（例如在带子上写一个0）。

然后图灵把H变了一下，将图灵机M也作为输入，计算H（M，M），记为H′。H′执行的步骤同H一样，它确定M处理它自身的编码M时会不会停机。不过，在得到结果后，H′执行的动作同H不一样。H在回答“是”或“否”后停机。H′则只有答案是“M对于编码M不会停机”时才会停机。如果答案是“M对于编码M会停机”，H′就进入死循环，永不停机。

这可能有点让人糊涂，希望你们能跟上。在计算理论的课程中，这是一个分水岭，一些学生会变得绝望，“我没法想明白这玩意！”一些学生则会拍手称快，“我喜欢这玩意！”

当初我也有些糊涂，不过我还是喜欢上了这些。也许你和我一样。我们来深呼吸一下，然后继续。

现在图灵抛出了最后的问题：

如果用H′自身的编码作为H′的输入，H′会怎么做呢？

到这个时候，拍手称快的学生也会开始头大。这确实很难想明白，不过我们还是试一试。

先假设H′对于输入H′不会停机。但这样就有个问题。前面说了，H′以图灵机M的编码作为输入时，如果M对于M会停机，H′就会进入死循环，反之则停机。因此，如果H′对于输入M不能停机，就意味着M对于输入M会停机。发现会有什么结果了吧？H′对于输入H′不会停机意味着H′对于输入H′会停机。但是H′不能既停机又不停机，这样就导致了矛盾。

因此假设H′对于H′不会停机是错误的，只能认为H′对于H′会停机。

但是这样又会有问题。只有在M对于M不会停机时，H′才会停机。因此如果H′对于H′不会停机，H′对于H′就会停机。又导致了矛盾。

这证明H′对于输入H′既不能停机也不能不停机。这是没法做到的，因此H′不可能存在。H′本身是H的特例，因此我们就证明了H不可能存在。

因此不存在明确程序能解决停机问题，这就是图灵的最后结论。停机问题证明了判定问题的答案是“否”；不存在明确程序能判定任意数学命题是否为真。图灵从而彻底埋葬了希尔伯特的这个问题。

从上面可以看出，图灵对停机问题不可计算性的证明，与哥德尔的不完备性定理具有同样的核心思想。哥德尔提出了可以编码数学命题的方法，从而让它们可以谈论自身。图灵则提出了编码图灵机的方法，让它们可以运行自身。

在这里我总结一下图灵里程碑式的成就。首先，他严格定义了“明确程序”的概念。其次，他提出的图灵机为电子计算机的发明奠定了基础。第三，他改变了大多数人的观念——计算存在局限。

哥德尔和图灵的命运

19世纪时，数学和科学被认为无所不能。希尔伯特和他的追随者认为他们即将实现莱布尼茨的梦想：发现自动判定命题的方法，并证明数学无所不能。类似的，在第2章我们看到，拉普拉斯相信，根据牛顿定律，科学家原则上能预测宇宙将发生的一切。

然而，20世纪早期在数学和物理上的发现表明，这个无所不能实际上并不存在。量子力学和混沌摧垮了精确预测的希望，哥德尔和图灵的结果则摧垮了数学和计算无所不能的希望。然而，图灵对停机问题的解决却为另一个伟大发现——可编程电子计算机——开辟了舞台。计算机后来给科学研究以及我们的生活带来了翻天覆地的变化。

在20世纪30年代发表他们的成果之后，图灵和哥德尔的命运迥异。同当时许多人一样，在希特勒和第三帝国出现后，他们的命运被彻底改变了。哥德尔受到时断时续的精神问题困扰，他在维也纳一直待到1940年，最后为了不被征入德军服兵役，移民到美国。（据他的传记作者王

浩说，[59] 在准备美国入籍面试时，他发现了美国宪法中的不一致性，结果他的朋友爱因斯坦在陪他去面试时只好不断同他聊天，以引开他的注意力。）

哥德尔和爱因斯坦一样，加入了声名卓著的普林斯顿高等研究院，并继续在数理逻辑领域做出重要贡献。然而，在20世纪60—70年代，他的精神状况不断恶化。去世前，他得了严重的妄想症，认为有人要毒害他。他拒绝进食，最终死于饥饿。

图灵也访问了普林斯顿高等研究院并得到了职位，但他决定回到英国。在第二次世界大战中，他加入了英国绝密的破解德军谜团密码（Enigma）的计划。以他的逻辑和统计学专长，再加上在电子计算上的成就，图灵领导研发了破译机器，最终几乎破解了所有使用谜团密码的情报。这使得英国在同德国作战时具有很大优势，并成为最终战胜纳粹的重要因素。

战后，图灵在曼彻斯特大学参与研制了第一批可编程电子计算机（基于通用图灵机的思想）。此后他的兴趣又回到探索大脑和身体的“计算”原理，他研究了神经学和生理学，并在发育生物学理论上做了有影响的工作，还探讨了智能计算机的可能性。然而他的生活与当时的社会道德习惯相抵触：他没有隐瞒自己的同性恋倾向。在20世纪50年代的英国同性恋是非法的，图灵因为与男性发生关系而被逮捕，并被判决接受药物“治疗”以改变他的“状况”。他也被取消了接触政府机密的权力。这些事件最终导致他在1954年自杀。有意思的是，哥德尔是因为怕被下毒而把自己饿死，图灵则是吃了有氰化物的苹果把自己毒死。图灵死的时候年仅41岁。

第5章　进化

一切伟大的真理开始时都是大逆不道。[60]

——萧伯纳（George Bernard Shaw），

《安纳扬斯卡，布尔什维克女皇》

（Annajanska, Te Bolshevik Empress ）

热力学第二定律认为封闭系统的熵会一直增加直至最大。这一点我们通过直觉就能认识到——不仅在科学上是这样，在我们的日常生活中，在我们对历史的理解中，在艺术、文学、宗教中都是这样。佛偈云，“一切行无常，生者必有尽。”[61] 旧约先知以赛亚预言“地必如衣服渐渐旧了”。[62] 莎士比亚则说：

哦，夏日的芳香怎能抵抗，[63]

多少个日夜前来猛烈地围攻，

就如岩石般顽强也无法坚守，

钢门结实，都得被时间磨空？

这真让人沮丧，一切都不可阻挡地走向最大熵。但是大自然中又能见到与之不符的例子：生命。不管从哪方面看，生命系统都是复杂的——它们处于有序和无序之间的某个地方。我们可以看到，在漫长的历史中，生命系统变得越来越精巧复杂，而不是熵逐渐增加，越来越无序。

我们已经知道，要让熵减少就必须做功。那么又是谁，是什么在创造和维持生命系统，并让它们越来越复杂呢？一些宗教认为这是神迹，但是在19世纪中叶，达尔文提出，生命进化是通过自然选择造就的。

还没有哪种科学思想像达尔文的进化论这样动摇过人类对于他们自身的观念。它也许是科学史上最具争议的思想，但它也是最好的思想。哲学家丹内特（Daniel Dennett）曾说：

如果要我选择一个历史上最重要的思想，[64] 我认为不是牛顿，也不是爱因斯坦，或是其他人，而是达尔文。自然选择的进化思想统一了生命和意义的疆域，还有可能会统一空间和时间、因果效应，机能和物理定律。

这一章将介绍达尔文进化论的历史和主要思想，以及进化如何产生出组织和适应。来自进化论的观念将一次又一次在这本书中出现。在第18章，我还将阐释这些观念如何被分子生物学以及复杂系统的研究成果进一步修正。

达尔文之前的进化观念

进化一词有“逐渐变化”的意思。生物进化就是生物的物种形态逐渐变化（有时候也很快）的过程。直到18世纪，物种形态都被认为是不会改变的；所有生物都是由神创造，从被创造出来就一直保持不变。在古希腊和印度曾有些哲学家认为人类可能是从其他物种变化而来，但是在西方，神创造万物的思想直到18世纪才开始被广泛质疑。

在18世纪中叶，达尔文提出进化论之前100年，法国动物学家布冯（George Louis Leclerc de Buffon）出版了鸿篇巨著《自然历史学》（*Historie Naturelle* ），书中描述了各种物种之间的相似之处。布冯认为地球的年龄远远大于圣经上说的6000年，而且现在所有物种都是由同一祖先进化而来，不过他没有说明进化的机制是怎样的。布冯在生物学和地质学上的研究对于之前的神创论是很大的突破。毫不奇怪法国天主教堂要将他的书烧掉。

查尔斯·达尔文的祖父厄拉斯莫斯·达尔文（Erasmus Darwin）也是一位杰出的学者，他也认为所有物种都是由同一祖先进化而来。他提出的进化机制是他孙子的自然选择理论的先驱。厄拉斯莫斯·达尔文在科学作品和诗中都表达了这种思想：

无垠波涛下的有机生命，[65]

在大海的呵护下孕育生长；

开始时很微小，显微镜下也看不见，

在淤泥中移动，在水中穿行；

随着一代一代繁衍，

逐渐获得了新的力量，具备了更强大的肢体；

从此开始出现数不清的植物群落，

和有鳍有脚有翅膀的会呼吸的动物。

比起现在的科学家来真是优雅多了！不过，和法国一样，英国天主教会也不喜欢这些思想。

达尔文之前最著名的进化论者是拉马克（Jean-Baptiste Lamarck，图5.1）。拉马克是法国贵族，也是一位植物学家，他在1809年出版了一本书——《动物哲学》（*Philosophie Zoologique* ），书中他提出了一种进化理论：新的物种从非生命物质中自发产生，然后物种会通过“获得性状的遗传”不断进化。这个思想认为生物在生命过程中会适应环境，而这种获得的适应性会直接遗传给后代。拉马克在书中举的一个例子是鹳类这样的涉水鸟类的长腿。他认为鹳类最初会不断伸展它们的腿，以免身体接触到水。经常伸展腿使得它们的腿越来越长，而这种获得的长腿性状又遗传给它们的后代，就这样越来越长。结果就造成了我们现在看见的涉水鸟类的长腿。



图5.1　拉马克（1744—1829）[图像出自《*LES CONTEMPORAINS N* 554：*Lamarck, naturaliste*（1744—1829）》，特莱（Louis Théret）著，博讷出版社（Bonne Press）1903年出版。Scientia Digital（http：//www.scientiadigital.com）拥有图像版权。经许可重印]

拉马克举了许多这样的例子。他还断言进化会产生“进步的趋势”，生物会进化得越来越高级，人类则是其巅峰。因此，生物的绝大部分变化都是越来越好，也越来越复杂。

当时绝大多数人都不接受拉马克的思想——不仅神创论者反对，相信进化的人也不认同。进化论者完全不信服拉马克对获得性状遗传的论证。事实上，他的经验数据也确实没有说服力，提到的生物特征的产生过程基本都只是他的臆测。

不过达尔文自己最初倒似乎很认同拉马克：“拉马克……有一些很确凿的证据，[66] 但是太过粗糙，他那些极具天赋的远见就好像科学中的先知、最自负的天才。”达尔文也认为，除了自然选择，获得性状的遗传也是进化的机制之一（这个思想在现在的“达尔文主义”中并没有留存下来）。

拉马克和达尔文都无法解释这样的遗传是如何发生的。在达尔文之后，随着遗传学的发展，才明确获得性状的遗传是不可能的。到20世纪初，拉马克的理论已没有什么影响，不过一些杰出的心理学家还是认为它能解释思维的某些方面，比如本能。弗洛伊德（Sigmund Freud）就表达了这种观点，“如果动物的本能生活能有任何解释，[67] 就只能是这个：它们将其经验赋予了它们的后代；也就是说，它们在记忆思维中保留了祖先的经验”。不过在弗洛伊德之后这些观念在心理学中也不再有影响。

达尔文理论的起源

达尔文（图5.2）是所有后进学生的榜样。孩童时期，他是才华横溢的家庭中成绩普通的学生。（他的父亲是一位成功的乡村医生，在达尔文十几岁时，有一次气得实在受不了，对他说：“除了打猎、养狗、抓老鼠，你还会什么？[68] 你只会让你自己和家族蒙羞！”）这个当时的后进学生后来却成了历史上最重要也最著名的生物学家。

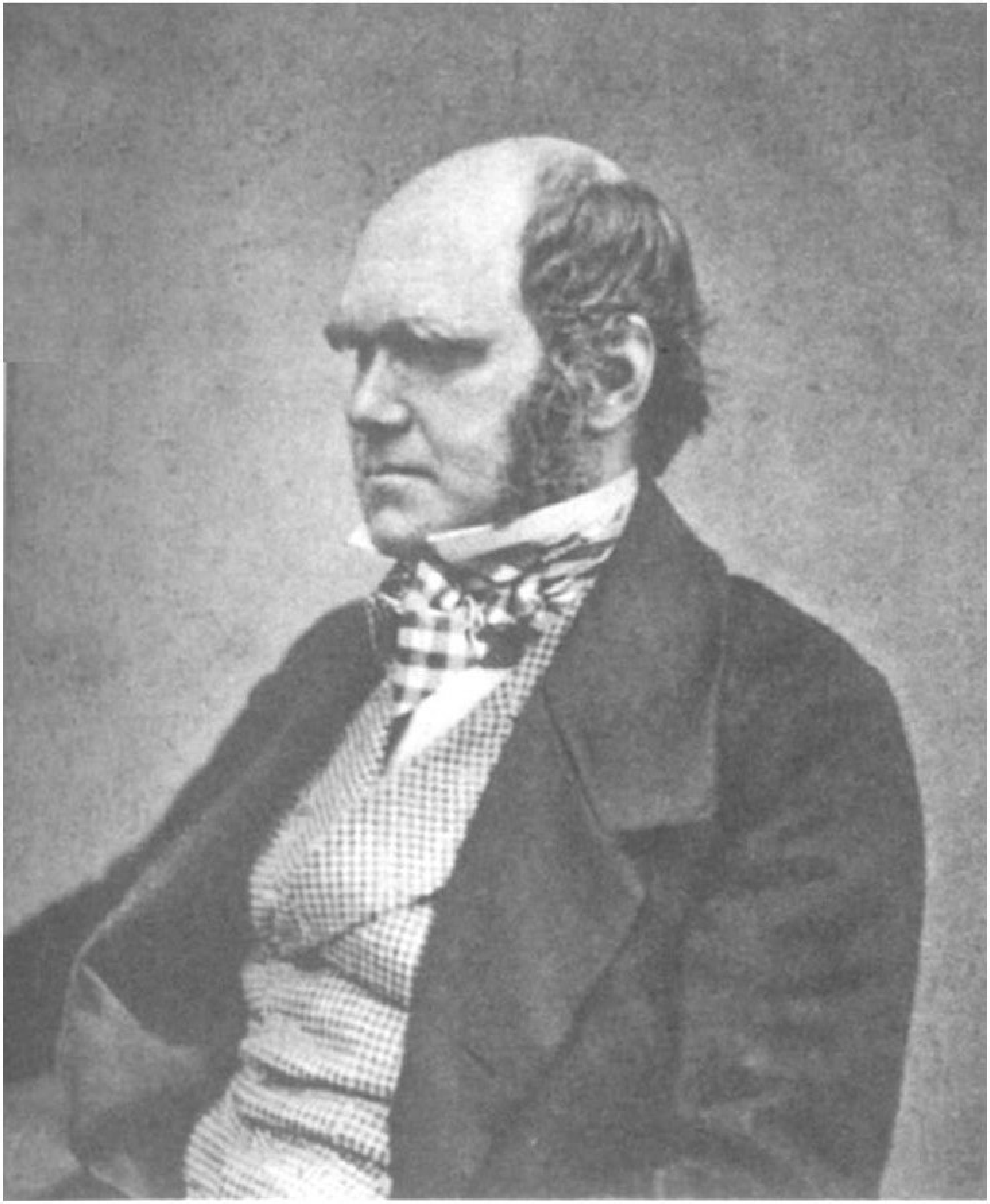


图5.2　达尔文（1809—1882），照片摄于1854年，数年后他出版了《物种起源》[照片来自范维尔（John van Wyhe）编辑的网上达尔文作品全集（http：//darwin-online.org.uk/），经许可引用]

1831年，在选择自己未来职业的时候（选择似乎包括乡村医生和乡村牧师），达尔文得到一个工作机会，在小猎犬号测绘船

（H.M.S.Beagle）上担任“博物学家”和“陪船长吃饭”。船长是一位“绅士”，旅途觉得寂寞，又不想和阶层低下的船员一起吃饭，就想找个绅士陪他吃饭。结果找到了达尔文。

达尔文在小猎犬号上待了将近5年（1831—1836），大部分时候在南非，除了陪船长吃饭，他还收集了许多植物、动物和化石标本，并且不断阅读、思考、写作。幸运的是他写了很多信，还保存了很多笔记，里面有很多观察、思考和阅读笔记。此后一生中他一直详细记录自己的思想。达尔文如果活到现在，肯定会热衷于写博客。

在随小猎犬号航行期间和之后，达尔文从多个学科的书籍和文章中汲取了大量思想。他信奉莱尔（Charles Lyell）的《地质学原理》（*Principles of Geology*，1830），认为各种地貌（山脉、峡谷和岩石的形成）是受风力、水流、火山爆发、地震等因素不断侵蚀而成，而不是圣经中所说的诺亚洪水这样的灾难造成的。这种渐进主义观点——微小因素日积月累也会有很大的影响——不容于当时的原教旨主义者，但是莱尔的证据让达尔文很信服，特别是航海经历让他看到了各种地质作用。

马尔萨斯（Thomas Malthus）的《人口学原理》（*Essay on the Principle of Population*，1798）让达尔文意识到群体数量的增长会导致对食物等资源的竞争。马尔萨斯论述的是人类数量的增长，达尔文却吸收其思想用来解释所有生物不断“为生存而斗争”，从而导致进化。

达尔文还读了亚当·斯密的自由市场圣经[69]——《国富论》（*The Wealth of Nations*，1776）。他通过这本书了解了斯密的经济的看不见的手的思想，大量个体只关心自己的私利，却使得整个社会的利益最大化。

通过在南非等地的观察，达尔文深深震惊于物种之间的巨大差异，以及不同物种对于环境的明显适应。他的一个最著名的例子是加拉帕格斯群岛（Galápagos）的燕雀，加拉帕格斯群岛距厄瓜多尔海岸约1000千米。达尔文发现燕雀的不同种群之间虽然很相似，它们的喙的大小和形状却差别很大。达尔文认识到岛上的不同燕雀都有共同的祖先，它们的祖先通过迁徙来到加拉帕格斯群岛。他还证明喙的形态是对不同食物来源的适应，[70]各小岛上的食物很不相同。达尔文认为各岛之间距离

遥远，加上环境差异很大，导致共同的祖先进化成了不同的物种。

可以想象，在达尔文航海期间和回到英国后，这些思想一直萦绕在他的脑海，他努力想要理解他看到的一切。微小的变化日积月累也会产生很大的影响。种群数量增长，资源却有限，因而不得不为生存而斗争。个体的自利行为却使得整体受益。生命的变化似乎有无限可能，各物种的性状似乎是针对它们所处的环境专门设计。物种从共同的祖先分化而来。

通过多年思考，这一切在他的脑中逐渐形成了统一的理论。在资源有限的情况下，生物个体存活的后代越多，则越占优势。后代不是完全复制父代，而是有一些细微的性状变化。有利于后代生存和繁衍的性状会被遗传给更多后代，从而在种群中扩散开来。慢慢地，通过繁殖时的随机变异和个体的生存斗争，就会形成适应环境的新物种。达尔文称这个过程为自然选择导致进化。

形成自己的理论之后数年里，达尔文只同莱尔等少数人交流了自己的思想。他之所以沉默，部分是由于他希望得到更多证据来支撑他的理论，更主要是因为他担心自己的理论会让宗教人士不快，尤其是他自己的夫人也是一位虔诚的宗教信徒。他再一次考虑成为乡村牧师，但又为自己的思想感到不安："（与我以前秉持的信念恰恰相反）我深信物种不是不可改变的[71]（这让我有沉重的负罪感）。"

另外，留下来的笔记还表明，达尔文当时已经意识到自己的理论对于人类地位的哲学意义。他写道："柏拉图……在《斐多篇》（Phaedo）中说[72]我们'与生俱来的思想'不可能来自经验，而是来自前世——但前世可能是猴子。"

竞争不仅是进化的中心要素，也是科学研究本身的主要动力。达尔文对发表成果的犹豫很快就消失了，因为他发现他可能会被抢先。1858年，达尔文收到了英国另一位自然学家华莱士（Alfred Russell Wallace）的手稿，《论变种无限地偏离原始类型的倾向》（*On the Tendency of Varieties to Depart Indefnitely from the Original Type*）。达尔文惊奇地发现华莱士也独立得出了自然选择导致进化的思想。达尔文在给莱尔的信中表达了自己的担心："我的所有成果，[73]不管意义有多大，也许都得不到承认。"然而，他还是慷慨地帮助华莱士发表了他的

论文，只是要求自己的成果也能同时发表，虽然他担心这个要求“有些可鄙”[74] 。

莱尔也认为，达尔文和华莱士应当同时发表他们的成果，以解决优先权问题。这个合作成果于1858年夏在林奈学会（Linnean Society）宣读。1859年底，达尔文出版了400多页的《物种起源》。

但优先权问题还是没有彻底解决。达尔文不知道，早在《物种起源》出版之前28年，一位不为人知的苏格兰人马修（Patrick Matthew）出版了一本标题和内容都很晦涩的书——《论海军木材和树木栽培》（*On Naval Timber and Arboriculture* ），书的附录中他提出了与达尔文的自然选择非常类似的思想。1860年，马修在杂志《加蒂纳记事》（*Gardiner’s Chronicle* ）上看到了达尔文的思想，就给杂志写了一封信申明他有优先权。达尔文心里也非常不安，他在信中回应道：“我完全承认马修先生多年前 [75] 就提出了我对于物种起源提出的自然选择解释……我只能向马修先生道歉，因为我完全不知道他的著作。”

那么自然选择的思想到底该归功于谁呢？显然，这又是一个同时独立发现的例子，一旦思想的时机成熟，就必然会有人想到。达尔文的同行赫胥黎（Thomas Huxley）就曾责骂自己：“真蠢，我怎么没有想到！”[76]

为什么最后是达尔文得到了一切荣誉呢？有几个原因，首先他当时已是声望很高的学者，最重要的是，与华莱士和马修的著作比起来，达尔文书中的思想更加清晰，给出的证据也多得多。是达尔文让自然选择从有趣而合理的猜测变成了极为完善的理论。

总结一下达尔文理论的主要思想：

◆存在进化，所有物种都来自共同的祖先。生命的历史就是物种呈树状分化。

◆一旦生物的数量超出了资源的承载能力，生物个体就会为资源竞争，从而导致自然选择。

◆生物性状会遗传变异。变异在某种意义上是随机的——变异并不必然会增加适应性（虽然前面提到达尔文自己接受拉马克的观点，认为

是这样的）。能够适应当前环境的变异更有可能被选择，也就是说具有这种变异的生物更有可能存活，并将这种新的性状遗传给后代，从而让后代中具有这种性状的个体增加。

◆进化是通过细微的有利变异不断累积逐渐形成的。

根据这个观点，自然选择导致的进化产物就像是被“设计”出来的，却不存在设计者。是机遇、自然选择和漫长的时间造就了这一切。熵的减少（生命系统结构越来越复杂，就像设计过的）是自然选择的结果。这个过程所需的能量来自生物从环境中获取的能量（阳光、食物等）。

孟德尔和遗传律

达尔文的理论没有解释性状如何从父代传给子代，也没有解释自然选择的基础——性状的变异——是如何产生的。直到20世纪40年代才发现DNA是遗传信息的载体。19世纪提出了许多遗传理论，但都没有产生很大影响，直到1900年，孟德尔（Gregor Mendel，图5.3）的工作被“重新发现”。

孟德尔是奥地利人，他是一位修道士，又是一位对自然有着强烈兴趣的物理教师。孟德尔在了解拉马克的获得性状遗传理论后，用豌豆做了一系列实验，以验证拉马克的理论，时间长达8年。他的结果不仅否定了拉马克的推测，同时也揭示了遗传的一些惊人的本质。

孟德尔研究了豌豆的几种性状：种子的光滑度和颜色；豆荚的形状；豆荚和花的颜色；花在植株上的位置以及植株的高度。每种性状都有两种不同的表现（例如，豆荚可以是绿色或黄色；植株可以是高或矮）。



图5.3　孟德尔（1822—1884）（引自国家医药图书馆，National Library of Medicine）（http：//wwwils.nlm.nih.gov/visibleproofs/galleries/technologies/dna.html）

直到现在，孟德尔的发现在遗传学中都被认为大致是正确的。首先，他发现植株的后代并不能遗传父代在生命期中获得的性状。因此拉马克式的遗传是不成立的。

另外他发现遗传是通过父母提供的离散“因子”产生的，每种性状对应父母提供的一种因子（也就是说，父母提供的一个因子决定了是高株还是矮株）。这里的因子大致对应于我们所说的基因。因此遗传的媒介是离散的，而不是像达尔文等人提出的是连续的（豌豆既可自花授粉也可异花授粉）。

孟德尔还发现，对于他研究的每一种性状，每一植株都有一对基因与之相对应（简单起见，我使用更为现代的术语，在孟德尔的时代并没有“基因”这个术语）。其中每个基因都对那种性状——例如高和矮——进行编码。这被称为等位基因（allele）。这样对于植株高度，其等位基因的编码就有三种可能：两者一样（高/高或矮/矮）或者不同（高/矮，与矮/高等同）。

不仅如此，孟德尔还发现，对于每一种性状，等位基因中有一个是显性的（例如高矮性状中高为显性性状），另一个则是隐性的（例如矮为隐性形状）。高/高个体总是表现为高株。高/矮个体也会表现为高株，因为高是显性的；只要有一个显性基因就够了。而只有矮/矮个体——两者都是隐性基因——才会表现为矮株。

举个例子，假设你用两株高/矮个体进行异花授粉。父母都很高，却还是有四分之一的可能他们的后代会从两者都遗传到矮基因，从而产生出矮/矮个体。

利用概率和推理，孟德尔能成功预测一代植株中表现出显性性状和隐性性状的植株各有多少。孟德尔的实验推翻了当时盛行的“混合遗传”的观念——认为子代的性状会是父母性状的平均。

孟德尔的研究是对遗传现象的第一个解释和量化预测，虽然孟德尔不知道他说的“因子”是什么构成的，也不知道它们如何通过交配重组。遗憾的是，1865年他的论文《植物杂交实验》发表在一个相当不著名的

期刊上，因而其重要性直到1900年才被承认，后来有几位科学家也通过实验得到了类似的结果。

现代综合

你可能会认为孟德尔的结果对达尔文主义会是极大的促进，因为它为遗传机制提供了实验验证。但其实在数十年里，孟德尔的思想都被认为是否定了达尔文的思想。达尔文的理论认为进化包括变异都是连续的（也就是说，生物个体之间的差异可以极为细微），而孟德尔的理论则提出变异是离散的（豌豆植株要么高要么矮，不能介于两者之间）。孟德尔理论的许多早期拥护者信奉突变学说（mutation theory）——认为生物变异是由于后代的突变，有可能非常大，并且自身产生进化，而自然选择只是用来保留（或消除）种群中这种突变的次要机制。达尔文及其早期追随者则坚决反对这种思想；达尔文理论的基石就是个体变异必须非常小，正是对这种微小变化的自然选择导致了进化，而且进化是渐进的。对于突变学说，达尔文有一句著名的驳斥，“Natura non facit saltum（自然不会跳跃）”。

达尔文主义者和孟德尔主义者相互论战了多年，直到20世纪20年代，人们发现，与孟德尔的豌豆的性状不同，生物的大部分性状都是由许多基因一起决定的，每个基因都有数个不同的等位基因，这种争论才烟消云散。多个不同等位基因会有数量极大的组合可能，从而使得生物的变异像是连续的。生物在基因层面的离散变异会导致表型——基因决定的生理特征（例如高矮、肤色等）——看似连续的变异。人们最终认识到，达尔文与孟德尔的理论并不矛盾，而是互补的。

早期达尔文主义者与孟德尔主义者之所以会水火不容，还有另一个原因，就是虽然双方都有实验证据支撑他们的立场，但当时却还没有成熟的概念体系（例如多个基因控制性状）和数学能将双方的理论融合到一起。要分析在杂交种群中多个基因在自然选择下相互作用的孟德尔式遗传的结果，必须发展出一套全新的数学工具。这套工具到20世纪20—30年代才由数学生物学家费希尔（Ronald Fisher）发展出来。

费希尔和高尔顿（Francis Galton）一起，创建了现代统计学。他最初是受现实世界中的农业和动物养殖问题的驱使。费希尔的成果，再加上霍尔丹（J.B.S.Haldane）和赖特（Sewall Wright）的工作，证明了达

尔文与孟德尔的理论实际上是一致的。不仅如此，费希尔、霍尔丹和赖特还提供了一个数学框架——群体遗传学（population genetics）——用来理解在孟德尔遗传学和自然选择作用下演化种群的等位基因的动力学。达尔文理论和孟德尔遗传学，[77] 再加上群体遗传学，共同形成了后来所谓的“现代综合（the Modern Synthesis）”。

费希尔、霍尔丹和赖特被认为是现代综合的奠基者。三人的意见有很多分歧，尤其是费希尔和赖特对自然选择和“随机基因漂移”的相对作用有激烈争议。在随机基因漂移过程中，某种等位基因占优势仅仅是因为随机的结果。例如，假设豌豆的高矮性状对植株整体的适应性没有影响。同时假设在某个时刻，仅仅是由于随机，群体中矮等位基因占的比重超过了高等位基因。如果每株高植株和矮植株产生的后代数量大致相同，则矮等位基因会更有可能在下一代中出现得更频繁，而这仅仅是因为具有矮等位基因的父代植株较多。通常，如果两种性状的选择优势没有差异，则其中一种性状最终会扩散至种群全体。漂移在小种群中的作用更强，因为在大种群中，漂移产生的微小波动会趋于被抹平。

赖特认为随机基因漂移在进化和新物种的产生中扮演了关键角色，而费希尔则认为漂移顶多是个次要角色。

双方的观点都有些道理，也很有趣。人们可能会认为，当英国人费希尔和美国人赖特碰面的时候，两人会一边喝啤酒，一边进行热烈而友善的讨论。然而两人富有成效的交流在他们各自发表文章攻击对方后结束了，到1934年，两人的通信基本终止。对于自然选择和随机漂移的相对作用的争论同以前孟德尔主义者与达尔文主义者之间的争议一样具有火药味——这真让人感到讽刺，因为正是费希尔和赖特的工作表明了双方的争议是不必要的。

现代综合在20世纪30—40年代得到了进一步发展，并形成了此后50年被生物学家普遍接受的一系列进化原则：

◆自然选择是进化和适应的主要机制。

◆进化是渐进过程，通过自然选择作用和个体非常细微的随机变异产生。这类变异在群体中大量发生，并且不存在偏好（也就是说并不是像拉马克认为的，必然会导致“进步”）。个体变异来源于随机基因突变

和重组。

◆宏观尺度上的现象，比如新物种的产生，可以用基因变异和自然选择的微观过程来解释。

现代综合的最初创建者认为他们解决了解释进化的主要问题，虽然他们还不知道基因的分子结构，也不知道变异产生的机制。就像进化学家塔特萨尔（Ian Tattersall）说的：“没有人在审视进化过程时能忽略现代综合的体系。[78] 这个体系不仅优雅有力，同时也让组织生物学各分支的研究者走到了一起，结束了纷争四起、互不认同、浪费精力的年代。”

对现代综合的挑战

对现代综合的合理性的严肃挑战始于20世纪60年代和70年代。最杰出的挑战者可能是古生物学家古尔德（Stephen Jay Gould，图5.4）和埃尔德雷奇（Niles Eldredge，图5.5），他们指出现代综合的预测与真实的化石记录并不相符。古尔德同时还是达尔文进化论最著名的鼓吹者和阐释者（他为普通读者写了大量书和文章），也是现代综合最激烈的批评者。

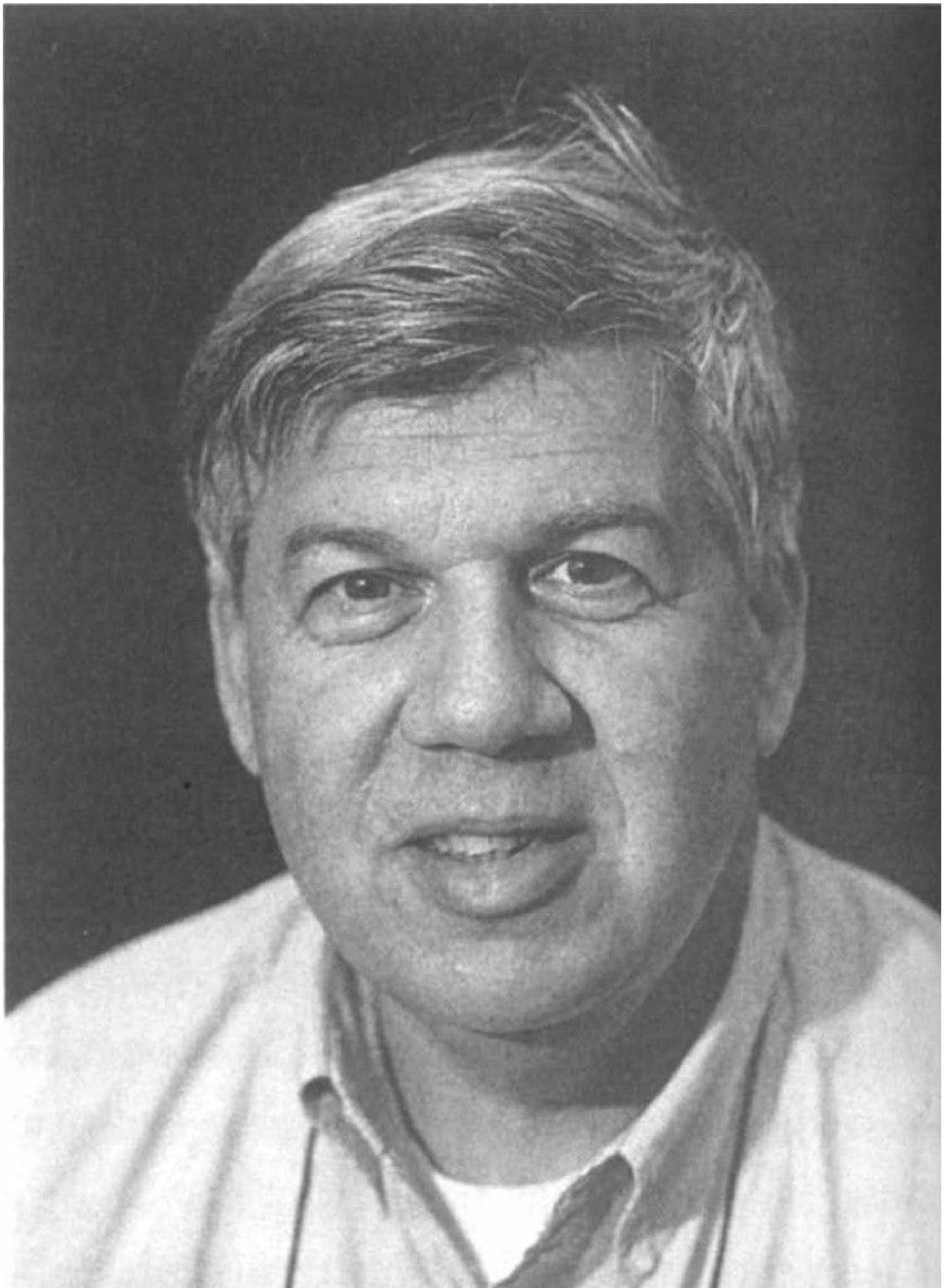


图5.4　古尔德（1941—2002）（Jon Chase/哈佛新闻办公室。哈佛大学版权所有，经许可可重印）

古尔德和埃尔德雷奇等人提出，现代综合预测的生物形态渐变不符合实际的化石记录：生物形态在很长时间里都没有变化（也没有新物种出现），而在（相对）很短的时间里形态却出现了剧烈变化，并产生出新的物种。这个特点被称为间断平衡（punctuated equilibria）。另有一些人则维护现代综合，认为化石记录很不完整，不能做出这样的推断。（一些诋毁间断平衡的人谑称其为“抽筋进化论”，古尔德则回击说渐进论拥护者支持的是“蠕动进化论”。）间断平衡在验证进化论的实验和进化计算机模拟中已被广泛观察到了。

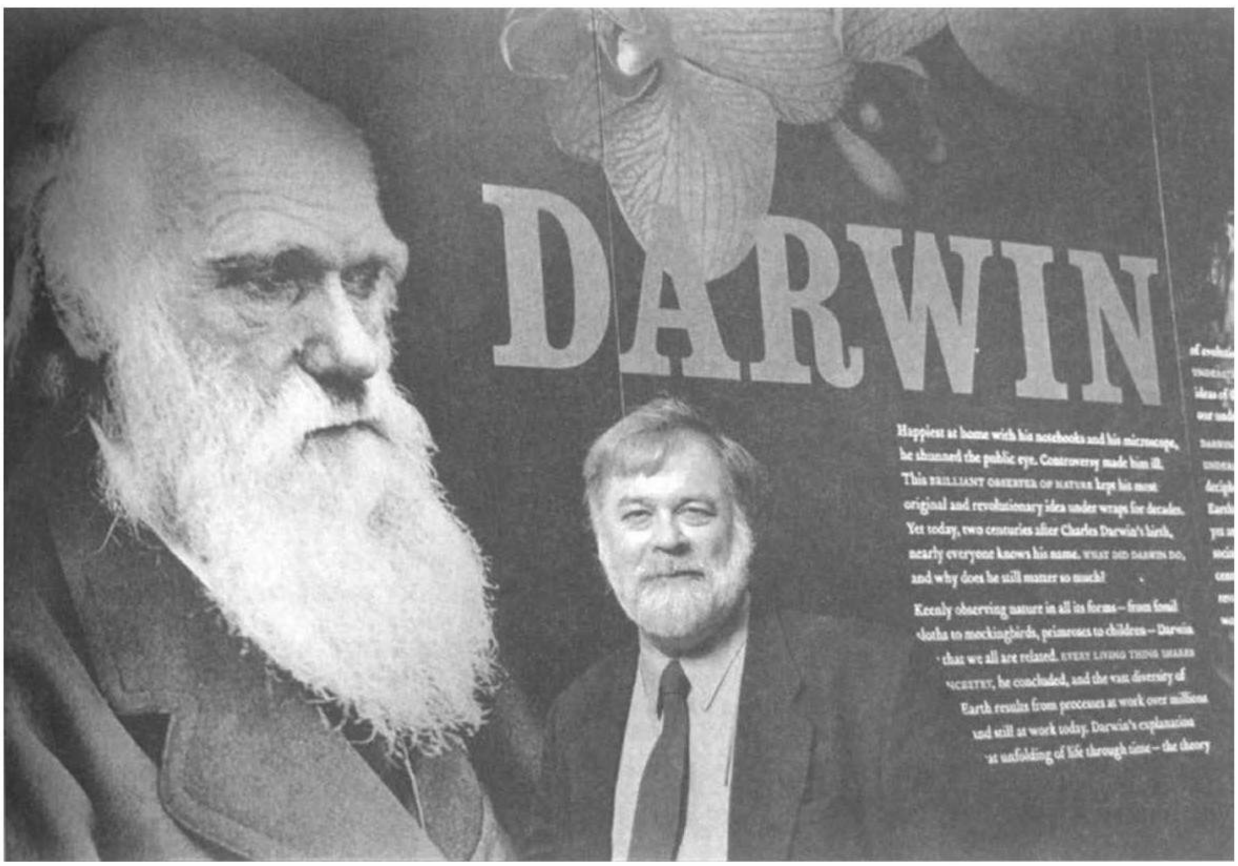


图5.5　埃尔德雷奇（照片由埃尔德雷奇本人提供）

因此古尔德与其合作者认为现代综合的“渐进主义”观点是错误的。他们还认为，其另外两个观点——自然选择和细微基因变异在生命史中起主要作用——也无法得到证据支持。

古尔德同意自然选择是进化很重要的机制，但他认为历史偶然和生物约束（biological constraints）的作用至少同样重要。历史偶然是指各种或大或小的随机事件都对生物的塑造有影响。一个例子就是流星的撞击摧毁了物种的栖息地，导致其灭绝，从而让新的物种得以产生。另一个例子就是未知的命运巧合让食肉哺乳动物比食肉鸟类更具优势，结果曾经很兴盛的食肉鸟类反而灭绝了。

对于偶然因素的作用，古尔德打了一个比方，想象一盘“生命录影带”，影带上记录了自地球诞生以来的一切进化事件。古尔德问，如果将影带倒回去重放，让初始条件稍微有些不同，又会怎么样呢？我们还会看到上次放映时进化出的类似生物吗？现代综合的回答可能会是“是”——自然选择仍然会修正生物以最好地适应环境，因此它们看上去会与实际发生的差不多。古尔德的回答是历史偶然会使得录影带重放时截然不同。

生物约束则是指自然选择所能创造的会有局限。显然自然选择不能违背物理定律——它不能创造出违反万有引力定律的飞禽或是无须进食的永动动物。古尔德等人认为，同物理约束一样，生物约束也对生物的进化有限制。

这个观点很自然延伸出一个结论，就是并不是生物的所有性状都能用“适应性”解释。饥饿感和性欲这些性状显然能增加我们的生存和繁衍机会。但有些性状可能是来源于偶然，或是适应性状和发育约束的旁效应。古尔德经常批评他所谓的“绝对适应论者”——他们坚持认为自然选择是复杂生物组织的唯一可能解释。

此外，古尔德等人还抨击了现代综合的第三个支柱，他们认为一些大尺度的进化现象无法用微观的基因变异过程和自然选择来解释。而需要自然选择作用于比基因和个体更高的层面——也许是整个种群。

古尔德质疑现代综合的一些证据来自分子进化。20世纪60年代，木村（Motoo Kimura）根据对蛋白质进化的观察提出了“中性进化”的理论，[79] 挑战自然选择在进化中的中心地位。20世纪70年代，化学家艾根（Manfred Eigen）和舒斯特（Peter Schuster）在RNA构成的病毒的进化中观察到了间断平衡现象，[80] 并发展出理论对其进行解释，认为进化的单位不是单个病毒，而是由原始病毒的变异复制体组成的病毒群

体，即准物种（quasi-species）。

进化论者根本不接受这些对现代综合的挑战，而且同达尔文主义以前的情况相似，争论经常充满敌意。1980年，古尔德写道，“综合理论事实上已经死了，[81] 虽然它还在教科书上被当做正统”。埃尔德雷奇和塔特萨尔甚至走得更远，他们声称，将进化归因于现代综合是“20世纪生物学最大的神话”。[82] 另一阵营中，杰出的生物学家麦尔（Ernst Mayr）和道金斯（Richard Dawkins）则坚决维护现代综合的信条。麦尔写道，“我认为进化综合的成果并没有严重错误，[83] 也不必被取代。”道金斯写道，“累积式的自然选择进化论[84] 是我们所知唯一能在原理上解释有组织复杂性的存在的理论。”现在许多人仍然持这个观点，然而，就像在第18章我们将看到的，随着新技术在遗传学中的应用，一些出人意料的发现已经深刻改变了人们对进化的认识，将自然选择导致的渐变作为塑造生命的唯一或主要力量的观点正受到越来越多的挑战。

必须强调的是，虽然古尔德和埃尔德雷奇等人挑战了现代综合的信条，但他们却同所有生物学家一样，仍然拥护达尔文主义的基本思想：进化在过去40亿年的生命史中一直存在，在未来也将继续存在；所有现代物种都是起源于共同的祖先；自然选择在进化中扮演了重要角色；并不存在“智能”引导了生物的进化或设计。

第6章　遗传学概要

近几十年来，分子生物学的进展已经改变了大多数生物学家对于进化的认识，并且对现代综合形成了挑战。

在第18章我将阐释其中一些结果，并介绍它们对遗传学和进化理论的影响。为了进行后面其他部分的讨论，这一章将简要介绍一些遗传学的基础知识。如果你对遗传学很熟悉，可以跳过这一章。

人们在19世纪早期发现所有生物都是由微小的细胞组成。在19世纪晚期又发现细胞的细胞核中有狭长的大分子，这种分子被称为染色体（因为它们很容易在实验中被染色），但是当时并不知道它们的功能。当时还发现细胞可以分裂成两个同样的细胞，从而复制自身，这个过程

被称为有丝分裂，分裂过程中染色体会进行复制。我们身体中许多细胞每过几小时就会分裂一次——这是身体发育、修复和日常维持的必要过程。

减数分裂也在同时期被发现，二倍体生物在产生卵子和精子时就是减数分裂。大部分哺乳动物都是二倍体生物，其他很多生物也是，它们的染色体（除了精子和卵子这些生殖细胞）都是成对出现（人类有23对）。减数分裂时，二倍体细胞会分裂成四个生殖细胞，每个生殖细胞的染色体数量为原细胞的一半。原细胞中每对染色体会断开，然后重组成新生殖细胞中的染色体。受精时，两个生殖细胞中的染色体会结合到一起，从而产生染色体数量正常的细胞。这样子代的染色体就是父母染色体的混合。这是有性生殖生物变异的主要来源。无性生殖的生物，后代与父代几乎一样。

这确实很复杂，也难怪生物学家们花了很长时间才弄清是怎么回事。但这还只是开始。

1902年，孟德尔的工作被重新发现后第二年，萨顿（Walter Sutton）首先提出染色体可能是遗传物质的载体。萨顿猜想染色体是由与孟德尔提出的因子相对应的单元（“基因”）组成，并且用减数分裂对孟德尔遗传学进行了解释。几年后，摩根（Thomas Hunt Morgan）用遗传学家的宠儿——果蝇——做实验，证实了萨顿的猜想。不过，基因的分子结构以及它们如何产生出生物的生理性状仍然是个谜。

到20世纪20年代末，化学家发现了核糖核酸（RNA）和脱氧核糖核酸（DNA），但是发现它们与基因的关联还要再过几年。接着又发现染色体中含有DNA，有些人开始怀疑DNA就是遗传物质。另一些人则认为细胞核中的蛋白质才是遗传物质。当然后来发现DNA才是正确答案，20世纪40年代中期最终通过实验证明了这一点。

但是还是有一些大的挑战。DNA到底是怎样决定生物性状的？比如说高株和矮株？细胞有丝分裂时DNA又是如何复制？作为自然选择基础的变异为何会发生在DNA层面上？

此后10年，这些问题都大致得到了解决。最大的突破发生在1953年，沃森（James Watson）和克里克（Francis Crick）发现，DNA的结构是双螺旋。20世纪60年代初，几位科学家一起成功破解了遗传密码

——DNA如何编码构成蛋白质的氨基酸。孟德尔无法知道基因的分子结构，但认识到了基因的存在，而现在基因终于可以用编码特定蛋白质的DNA片段进行定义了。很快又发现了编码如何被细胞转化为蛋白质，DNA如何复制自身，以及复制错误、突变和性重组如何引起变异。遗传学研究从此被引爆，此后迅速发展，直至现在。

生物的所有性状——表型——几乎都是源自细胞中蛋白质的特性及其相互作用。蛋白质是由氨基酸组成的长链分子。

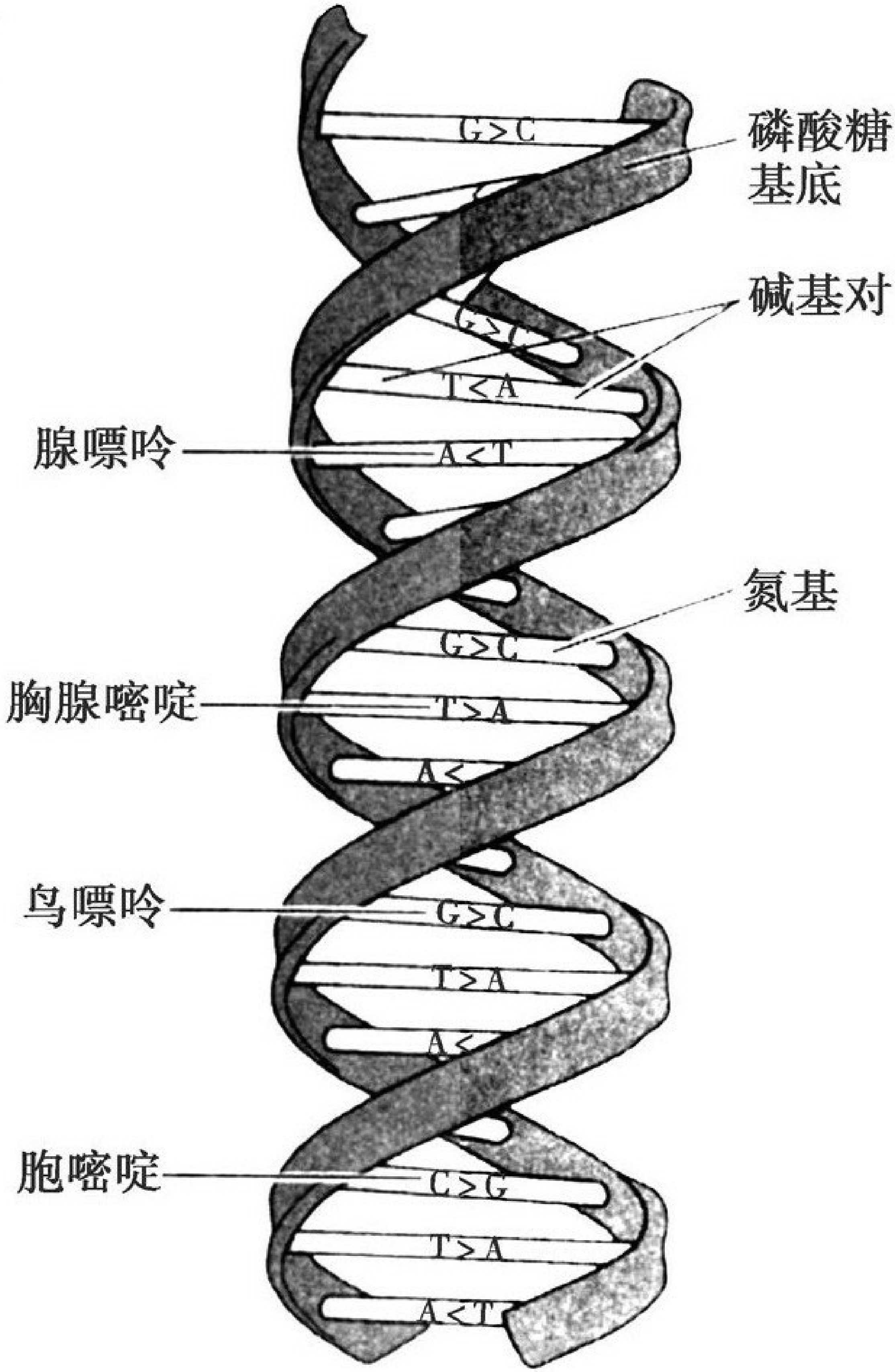
你身体中每个细胞都有几乎一样的完整DNA序列，DNA序列由核苷酸连在一起组成。核苷酸含有称为碱基的化合物，碱基有四种形式，（缩写为）A、C、G、T。人类的DNA序列实际上是由A、C、G、T分子对组成的双线。化学势使得A总是与T配对，C则与G配对。

序列经常用两行符号（碱基对）表示，例如：

T C C G A T T……

A G G C T A A……

在DNA分子中，双线相互缠绕，形成一条双螺旋（图6.1）。



磷酸糖
基底
G>C
碱基对
G>C
T<A
腺嘌呤
A<T
氮基
G>C
胸腺嘧啶
T>A
鸟嘌呤
G>C
T>A
胞嘧啶
C>G
T>A
A<T

图6.1　DNA的双螺旋结构[美国国家人类基因组研究所，基因研究组语音词汇（http：//www.genome.gov/glossary.cfm.）]

基因就是由DNA的序列片段组成。大致上，一个基因对应于一个特定的蛋白质。基因编码了构成蛋白质的氨基酸。氨基酸的编码方式就是遗传密码。这个编码对地球上的所有生物几乎都是一样的。三个碱基对应一种氨基酸。例如AAG就对应苯基丙氨酸，CAC则对应缬氨酸。这种三联体被称为密码子。

基因又是如何构造蛋白质的呢？每个细胞都有一套复杂的分子机制来进行这个工作。第一步是转录（图6.2），这一步是在细胞核中进行。一种被称为核糖核酸聚合酶的活性蛋白会从双螺旋的一边松开一小段DNA。然后这种酶会用DNA的一边产生出信使RNA分子（m RNA），m RNA逐字复制DNA片段。实际上是反拷贝：如果基因上为C，m RNA上则对应为G，如果基因上为A，m RNA上就对应为U（m RNA版本的T）。通过反复制又可以重构出原来的序列。

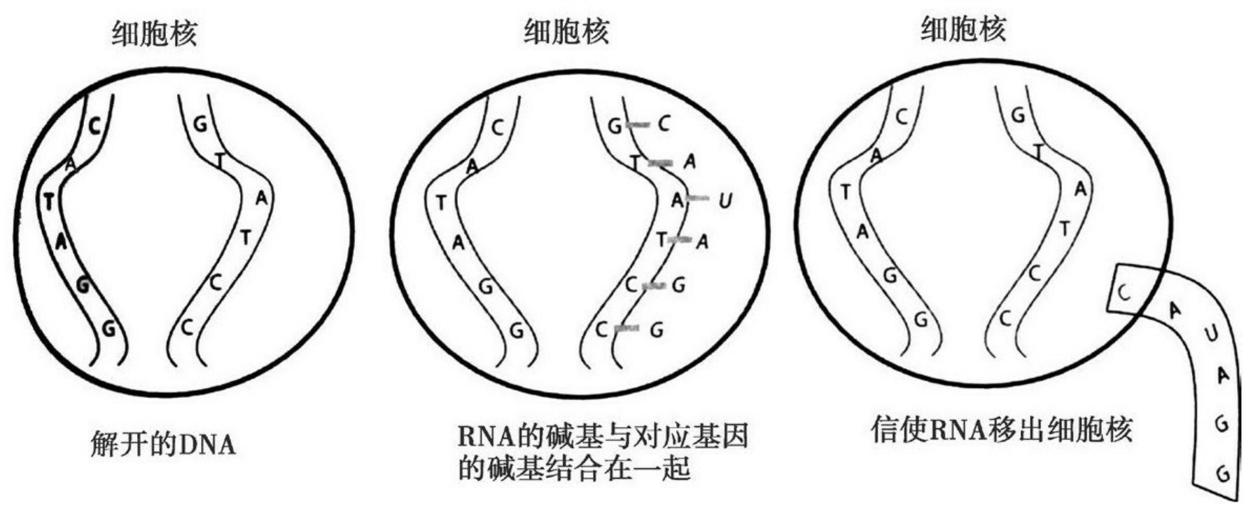


图6.2　DNA转录为信使RNA。注意DNA中是符号T, RNA中则是符号U

这个转录过程一直持续到基因完全转录成m RNA。

第二步是翻译（图6.3），这是在细胞质中进行。新产生的m RNA序列从细胞核进入细胞质，在这里细胞质结构核糖体将m RNA上的密码子逐个读出。在核糖体中，各个密码子会与转运RNA分子（t RNA）上的反密码子结合。反密码子是由互补碱基组成。例如，在图6.3中，被转录的m RNA密码子是U A G，反密码子则是互补碱基A U C。如图6.3

所示，t RNA分子上的反密码子会与m RNA的密码子相连。而t RNA分子的反密码子则会组合成相应的氨基酸（密码子A U C实际就是异亮氨酸的编码）。侯世达将t RNA比喻为“细胞的闪存卡”。[85]

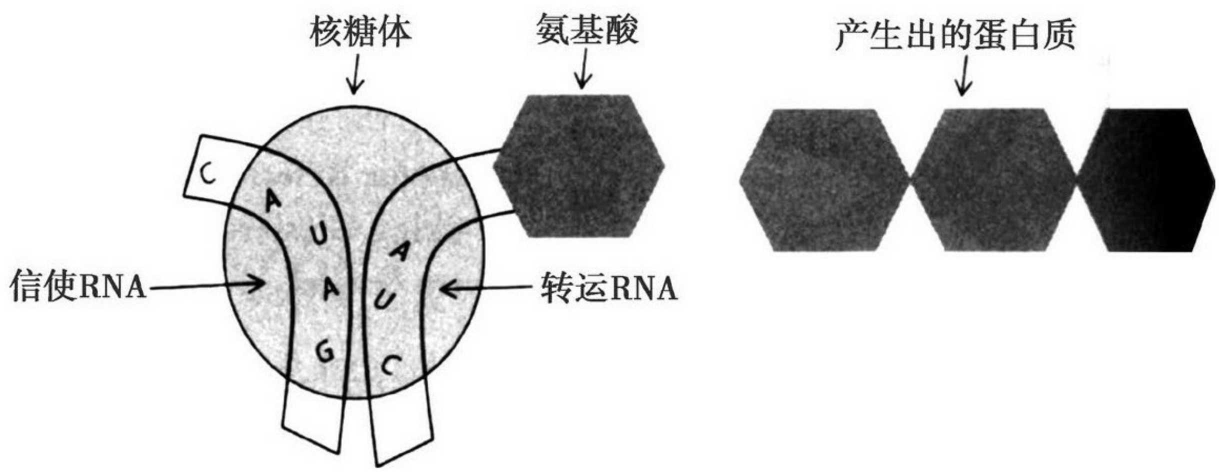


图6.3 信使RNA转译成氨基酸

核糖体将氨基酸从t RNA分子上分离下来并将它们合成为蛋白质。一旦遇到终止密码子，核糖体就会收到停止信号，然后将蛋白质释放到细胞质，让它们去执行自己的功能。

基因的转录和翻译就称为基因表达。

所有这一切都在亿万个细胞中不断进行。神奇的是这一切所需的能量非常少——如果你坐着看电视，所有亚细胞层面的活动每小时消耗的能量不会超过418焦。这是因为这些过程依靠的是分子的随机运动和大量的碰撞，只需从“环境热源”（比如你温暖的房间）中获取能量就够了。

碱基的配对特性，A配T， C配G，也是DNA复制的关键。在有丝分裂开始时，酶会将DNA的双螺旋解开。然后其他酶会读取两条DNA上的核苷酸，并将新的核苷酸附到上面（在细胞中新的核苷酸会不断被制造出来），A连到T， C连到G。这样，DNA就被复制成了两个新的DNA双螺旋，每个新细胞都得到一份完整的DNA拷贝。细胞中有许多机制保证复制正常进行，但是偶尔也会发生错误（碱基配对错误，大约1000亿个核苷酸产生一次），从而导致变异。

值得注意的是，这其中含有绝妙的自指特性：所有这些决定DNA的

转录、翻译和复制的复杂细胞机制——m RNA、t RNA、核糖体、聚合酶，等等——本身都编码在DNA中。就像侯世达说的：“DNA中包含其本身的解码者的编码！”它也包含合成核苷酸的所有蛋白质的编码，而核苷酸是构造DNA的材料。如果图灵还活着，看到这种自指特性肯定会非常高兴。

20世纪60年代中期，遗传学家们为研究这个极度复杂的系统进行了卓绝的努力，最终理解了上面这些基本过程。这些努力也导致了对分子层面进化的新理解。

1962年，克里克、沃森和生物学家威尔金斯（Maurice Wilkins）因为揭示了DNA的结构而分享了诺贝尔生物与医学奖。1968年，科拉纳（Har Gobind Korana）、霍利（Robert Holley）和尼伦伯格（Marshall Nirenberg）因为破解了遗传密码而获得了同一奖项。至此，进化和遗传的主要秘密似乎基本都被发现了。然而，在第18章我们将看到，它的复杂程度实际上远远超出了所有人的想象。

第**7**章　度量复杂性

这本书讲的是复杂性。但是到现在书中还没有严格定义这个术语，也没有明确回答以下问题：人类大脑比蚂蚁的大脑复杂吗？人类基因组是不是比酵母菌的基因组复杂？生物的复杂性在进化过程中是不断增加吗？直观上这些问题的答案太明显不过了。然而，要想得出一个公认的复杂性定义，来回答这些问题，其中的困难却超乎想象。

2004年我曾在圣塔菲复杂系统暑期班上组织过一个研究小组。那一年有点特别，因为是圣塔菲研究所创建20周年。小组中有圣塔菲一些最杰出的学者，包括法墨尔（Doyne Farmer）、克鲁奇菲尔德（Jim Crutchfield）、弗瑞斯特（Stephanie Forrest）、史密斯（Eric Smith）、米勒（John Miller）、胡伯勒（Alfred Hübler）和艾森斯坦（Bob Eisenstein）——都是物理、计算机、生物、经济和决策论等领域的知名学者。暑期班的学生——研究生和博士后层次的青年科学家——在讨论班上可以提问。第一个问题就是：“复杂性该怎样定义？”听到后大家都笑了起来，因为这个问题是如此直截了当，如此让人期待，然而又是如此难以回答。然后多位学者对这个术语给出了各种不同的定义，接着彼

此之间又产生了一些争论。学生们都一头雾水。就连圣塔菲这个复杂系统领域最著名研究所的学者对复杂性的定义都达不成共识，复杂性科学又是如何产生的呢？

答案是复杂性科学不止一个，而是有好几个，每个对复杂性的定义都不一样。其中一些定义很正式，一些则不那么正式。如果想要有统一的复杂性科学，就得弄清楚这些正式或非正式概念之间的关联。要对过于复杂的复杂性概念进行尽可能的提炼。这项工作目前还远未结束，也许还要等待那些被搞得一头雾水的下一代科学家来完成。

我希望那些同学不要对此太过诧异。只要了解一点科学史就能明白，核心概念缺乏公认的定义是很普遍的。牛顿对力的概念就没有很好的定义，事实上他不是很喜欢这个概念，因为它需要一种魔术般的“远距离作用”，而这在对自然的机械论解释中是不允许的。遗传学作为生物学领域发展最快和最大的学科，对于如何在分子层面上定义基因的概念[86]也没有达成一致。天文学家发现宇宙95%都是由暗物质和暗能量组成，却不清楚暗物质和暗能量到底是什么。心理学家对思维和概念也没有明确的定义，更不知道它们在大脑中对应的是什么。这还只是部分例子。科学的进步往往就是通过为尚未完全理解的现象发明新术语实现的：随着科学逐渐成熟，现象逐渐被理解，这些术语也逐渐被提炼清晰。例如，物理学家现在就理解了自然界中所有的力都是四种基本力的组合：电磁力、强相互作用、弱相互作用、引力。基本粒子“远距离作用”的现象也已经被理论化。在量子力学中发展出描述四种基本力的统一理论是物理学现在面临的最大挑战。也许将来我们也会将“复杂性”分解成几个基本方面，并最终将这几个方面结合起来，形成对复杂现象的全面理解。

2001年，物理学家劳埃德（Seth Lloyd）发表了一篇文章，[87]提出了度量一个事物或过程的复杂性的三个维度：

描述它有多困难？

产生它有多困难？

其组织程度如何？

劳埃德列出了40种度量复杂性的方法，这些方法分别是从动力学、

热力学、信息论和计算等方面来考虑这三个问题。我们已经了解了这些概念的背景，现在我们可以来看看其中一些定义。下面我会通过比较人类基因组与酵母菌基因组的复杂性来阐释这些定义。人类基因组大约有30亿组碱基对（即核苷酸对）。据估计人类大约有25000个基因——也就是对蛋白质进行编码的区域。让人吃惊的是，只有2%的碱基对组成了基因；其余的非基因部分被称为非编码区。非编码区有几个功能：其中一些用来防止染色体解体；一些则帮助调控真正基因的运作；有一些则可能是没有任何作用的“垃圾”或者功能还没有被发现。

你肯定听说过人类基因组计划，但你可能不知道还有一个酵母菌基因组计划，这个计划的目标是测定几种酵母菌的完整DNA序列。测出的第一种被发现大约有1200万组碱基对和6000个基因。

用大小度量复杂性

复杂性的一个简单度量就是大小。根据这个度量，如果比较碱基对数量，人类比酵母复杂250倍，如果比较基因数量，人类则只比酵母复杂4倍。

250倍还是蛮多的，看来人类还是挺复杂，至少比酵母复杂。不过单细胞变形虫的碱基对是人类的225倍，拟南芥的基因与人类的大致一样多。

人类显然要比变形虫或芥菜复杂，至少我希望是这样。这就表明用基因组的规模来度量复杂性并不合适；我们的复杂性应该是某种比碱基对或基因的绝对数量更深刻的东西（图7.1）。

用熵度量复杂性

另一种直接的复杂性度量就是香农熵，在第3章曾将香农熵定义为信息源相对于信息接收者的平均信息量或“惊奇度”。举个例子，假设消息由符号A、C、G和T组成。如果序列高度有序，很容易描述，例如“A A A A A A A……A”，则熵为零。完全随机的序列则有最大可能熵。

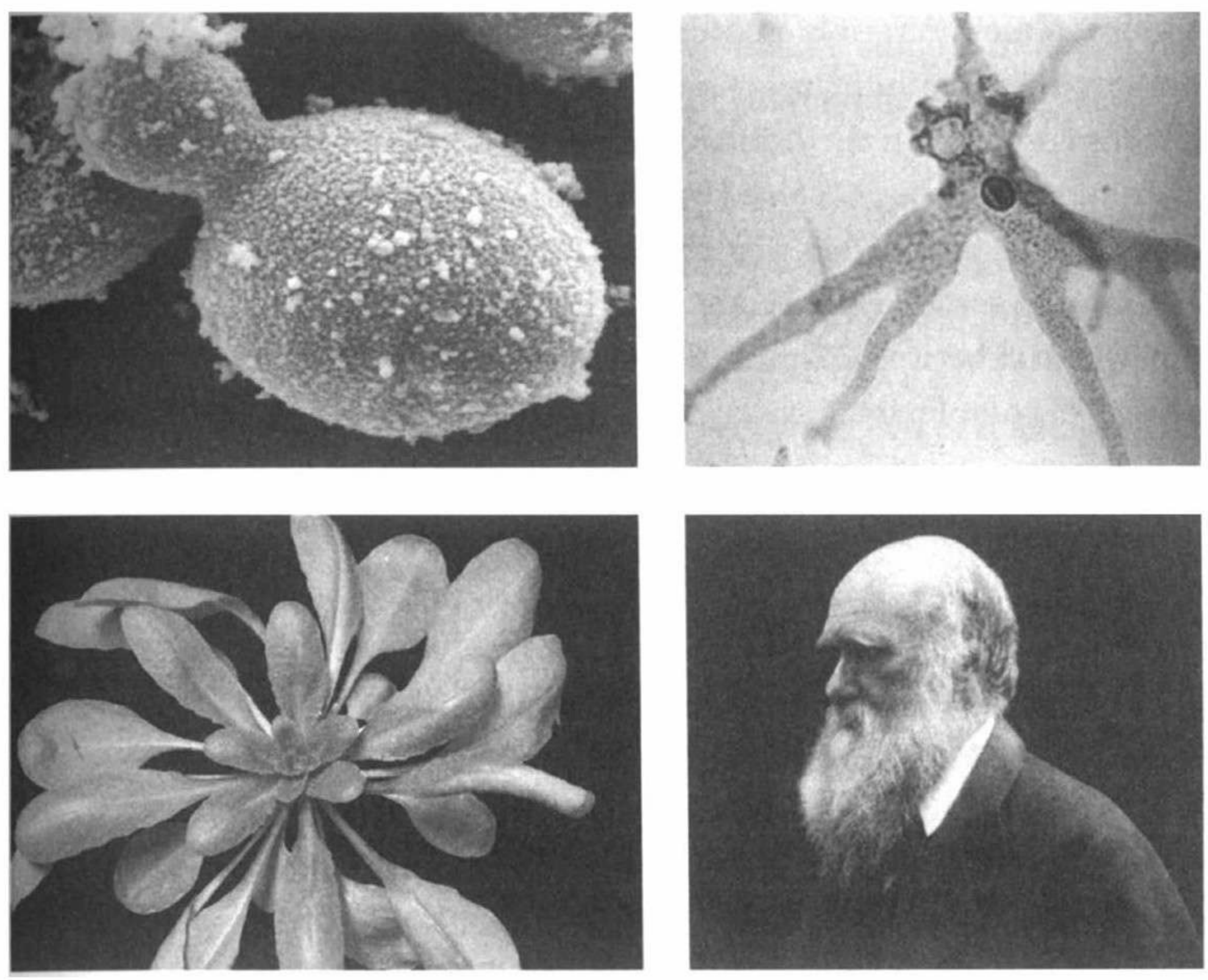


图7.1从左上角依顺时针分别是：酵母、变形虫、人类、拟南芥。哪个最复杂？如果用基因组长度度量复杂性，那变形虫毫无疑问会跑冠军（如果它有腿的话）。[酵母照片来自NASA（http：//www.nasa.gov/mission_pages/station/science/experiments/Yeast-GAP.html）；变形虫照片来自NASA（http：//ares.jsc.nasa.gov/astrobiology/biomarkers/_images/amoeba.jpg）；拟南芥照片由Kirsten Bomblies提供；人类照片来自范维尔（John van Wyhe）编辑的网上达尔文作品全集（http：//darwin-online.org.uk/），经许可引用]

用香农熵度量复杂性有一些问题。首先，所针对的对象或过程必须像上面一样转换成某种“消息”的形式。这并不总是那么容易做到，例如，人类大脑的熵该怎么度量呢？另外，随机消息的熵最高。我们可以随机排列A、C、G和T来人工构造一个基因组，这个随机的基因组几乎不可能有用，却会被认为比人类基因组更复杂。很显然，正是因为基因组不是随机的，而是不断进化从而让基因更有利于我们的生存，例如控制我们的眼睛和肌肉发育，才使得人类如此复杂。最复杂的对象不是最有序的或最随机的，而是介于两者之间。简单的香农熵不足以抓住我们对复杂性的直观认识。

用算法信息量度量复杂性

人们提出了许多改进方法来用熵度量复杂性。其中最著名的方法由柯尔莫哥洛夫（Andrey Kolmogorov）、查汀（Gregory Chaitin）和索罗蒙洛夫（Ray Solomonoff）分别独立提出，他们将事物的复杂性定义为能够产生对事物完整描述的最短计算机程序的长度。这被称为事物的算法信息量。[88]“例如，考虑一个很短的（人工）DNA序列：

A C A C A C A C A C A C A C A C A C A C（序列1）

一个很短的计算机程序，“打印A C 10次”，就能输出这个序列。因此这个序列的算法复杂度很低。作为比较，下面是我用伪随机数发生器生成的一个序列：

A T C T G T C A A G A C G G A A C A T（序列2）

如果我的随机数发生器没有问题，这个序列就不会有可识别的特征，因此程序要长一些，比如“打印字符串A T C T G T C A A A A C G G A A C A T”。显然序列1可以压缩，而序列2则不能，因而包含更多算法信息。与熵类似，随机对象的算法信息量也会比我们直观上认为复杂的事物的信息量更大。

物理学家盖尔曼（Murray Gell-Mann）提出了一种称为“有效复杂性（effective complexity）”的相关度量，[89] 更符合我们对复杂性的直观认识。盖尔曼认为任何事物都是规则性和随机性的组合。例如，序列1就有非常简单的规则性：重复的A C模式。序列2则没有规则性，因为它是随机产生的。与之相比，生物的DNA则有一些规则性（例如，基因组不同部分之间存在重要关联），也有一些随机性（例如DNA中的垃圾）。

为了计算有效复杂性，首先要给出事物规则性的最佳描述；有效复杂性定义为包含在描述中的信息量或规则集合的算法信息量，两者等价。

序列1具有规则性，即A C不断重复。描述这个规则性所需的信息量就是它的算法信息量：程序“打印A　C数次”的长度。因此，事物的结构可预测性越大，有效复杂性就越低。

序列2处于另一个极端，因为是随机的，所以没有规则性。因而也不需要信息来描述，虽然序列本身的算法信息量是最大的，序列规则性的算法信息量——其有效复杂性——却为零。简而言之，就如我们希望的，最有序和最随机的事物有效复杂性很低。

能够发育的生物的DNA具有许多独立和相关的规则性，这些规则需要可观的信息才能描述，因此会有很高的有效复杂性。显然，问题是我们如何给出这些规则？如果不同观察者对于系统的规则不能达成一致又怎么办？

盖尔曼用科学理论的形成做了类比，科学理论的形成实际就是寻找自然现象规律的过程。对于任何现象，都有多个描述其规律的可能理论，但显然理论有好有差，一些更加简洁优雅。盖尔曼对这个很有经验，他极为优雅的理论厘清了当时让人混淆的基本粒子类型及其相互作用，这让他获得了1969年的诺贝尔物理学奖。

类似的，对于提出的一个事物的各种不同规则集，我们可以利用奥卡姆剃刀（Occam’s Razor）来决定哪个是最好的。最好的规则集是能描述事物的最小规则集，同时还能将事物的随机成分最小化。例如，生物学家们现在已经发现了人类基因组的许多规律，包括基因、基因之间的相互作用，等等。但还有许多似乎不遵循任何规则的随机方面——也就是所谓的垃圾DNA。如果生物学的盖尔曼出现，他也许会找到约束更多基因组的极为简单的规则集。

有效复杂性是很有吸引力的思想，虽然同其他许多度量复杂性的提议一样，很难实际操作。也有批评意见指出，这个定义中的主观性也有待解决。[90]

用逻辑深度度量复杂性

为了更加接近我们对复杂性的直觉，数学家班尼特在20世纪80年代初提出了逻辑深度（logical depth）的概念。一个事物的逻辑深度是对构造这个事物的困难程度的度量。高度有序的A、C、G、T序列（例如前面的序列1）显然很容易构造。同样，如果我要你给我一个A、C、G、T的随机序列，你也很容易就可以做出来，用个硬币或骰子就可以了。但如果我要你给我一个能够生成可发育的生物的DNA序列，如果不偷看

真正的基因组序列，别说你，任何一个生物学家都会觉得很难办到。

用班尼特的话说，“有逻辑深度的事物[91]……从根本上必须是长时间计算或漫长动力过程的产物，否则就不可能产生”。或是像劳埃德说的，“用最合理的方法生成某个事物时需要处理的信息量[92]等同于这个事物的复杂性，这是一个很吸引人的想法”。

为了更精确地定义逻辑深度，班尼特将对事物的构造换成了对编码事物的0/1序列的计算。例如，我们可以用两位二进制数来编码核苷酸符号：A=00，C=01，G=10，T=11。用这个编码，我们就能将A、C、G、T转换成0/1序列。然后编写一个图灵机，用编写好的图灵机在空白带子上产生出这个序列，所需要的时间步就是其逻辑深度。

一般而言，多个图灵机都能产生出这个序列，所需的时间也可能不一样多。班尼特还必须说明应该用哪一个图灵机。他提出，应该根据前面提到的奥卡姆剃刀原则，选最短的那个（也就是状态和规则最少的那个）。

逻辑深度具有很好的理论特征，符合我们的直觉，但是也没有具体给出度量实际事物复杂性的方法，因为没有寻找生成指定事物的最小图灵机的可操作方法，更不要说如何确定机器运算所需的时间。此外，也没有考虑将事物表示成0/1序列的困难。

用热力学深度度量复杂性

20世纪80年代末，劳埃德和裴杰斯（Heinz Pagels）提出了一种新的复杂性度量[93]——热力学深度（thermodynamic depth）。劳埃德和裴杰斯的思想与班尼特的思想很相似：越复杂的事物越难构造。不过与图灵机生成对事物的描述所需的时间步不同，热力学深度首先是确定“产生出这个事物最科学合理的确定事件序列”，然后测量“物理构造过程所需的热力源和信息源的总量”。[94]

例如，要确定人类基因组的热力学深度，我们得从最早出现的第一个生物的基因组开始，列出直到现代人类出现的所有遗传演化事件（随机变异、重组、基因复制等等）。可以想象，人类进化出来的时间比变形虫要长10亿年，热力学深度肯定也大得多。

同逻辑深度一样，热力学深度也只是在理论上有意义，要真的用来度量复杂性也存在一些问题。首先，我们要能列出事物产生过程中的所有事件。另外，也有批评意见指出，[95] 劳埃德和裴杰斯的定义中没有明确界定什么是“事件”。一次遗传变异到底是单个事件还是在原子和亚原子层面导致变异发生的上百万次事件呢？两个祖先的基因重组应当视为单个事件吗？还是应当将导致它们相遇、交配和产生后代的所有微观事件都包括进来呢？用更专业一点的话说，是不清楚如何将系统的状态“粗粒化”——也就是说，在列出事件时，如何确定哪些是相关的宏观状态。

用计算能力度量复杂性

如果复杂系统能够执行计算，不管系统是天然的还是人工的，也许有可能用它们的计算的复杂程度来度量它们的复杂性。像物理学家沃尔夫勒姆（Stephen Wolfram）就提出，[96] 系统的计算能力如果等价于通用图灵机的计算能力，就是复杂系统。不过，班尼特等人则认为，[97] 具有执行通用计算的能力并不意味着系统本身就是复杂的；我们应当测量的是系统处理输入时的行为的复杂性。譬如，通用图灵机本身并不复杂，但是有了程序和输入，进行了繁复的计算，就能产生复杂的行为。

统计复杂性

物理学家克鲁奇菲尔德和卡尔·杨（Karl Young）定义了一个称为统计复杂性[98]（statistical complexity）的量，度量用来预测系统将来的统计行为所需的系统过去行为的最小信息量。[物理学家格拉斯伯杰（Peter Grassberger）也独立给出了很类似的定义，称为有效度量复杂性。]统计复杂性与香农熵相关，定义中系统被视为“消息源”，其行为以某种方式量化为离散的“消息”。对统计行为的预测需要观测系统产生的信息，然后根据信息构造系统的模型，从而让模型的行为在统计上与系统本身的行为一致。

例如，序列1的信息源模型可以很简单：“重复A C”；因此其统计复杂性很低。然而，与熵或算法信息量不同，对于产生序列2的信息源也可以有很简单的模型：“随机选择A、C、G或T。”这是因为统计复杂性模型允许包含随机选择。统计复杂性的度量值是预测系统行为的最简单模型的信息量。因此，与有效复杂性一样，对于高度有序和随机的系

统，统计复杂性的值都很低，介于两者之间的系统则具有高复杂性，与我们的直觉相符。

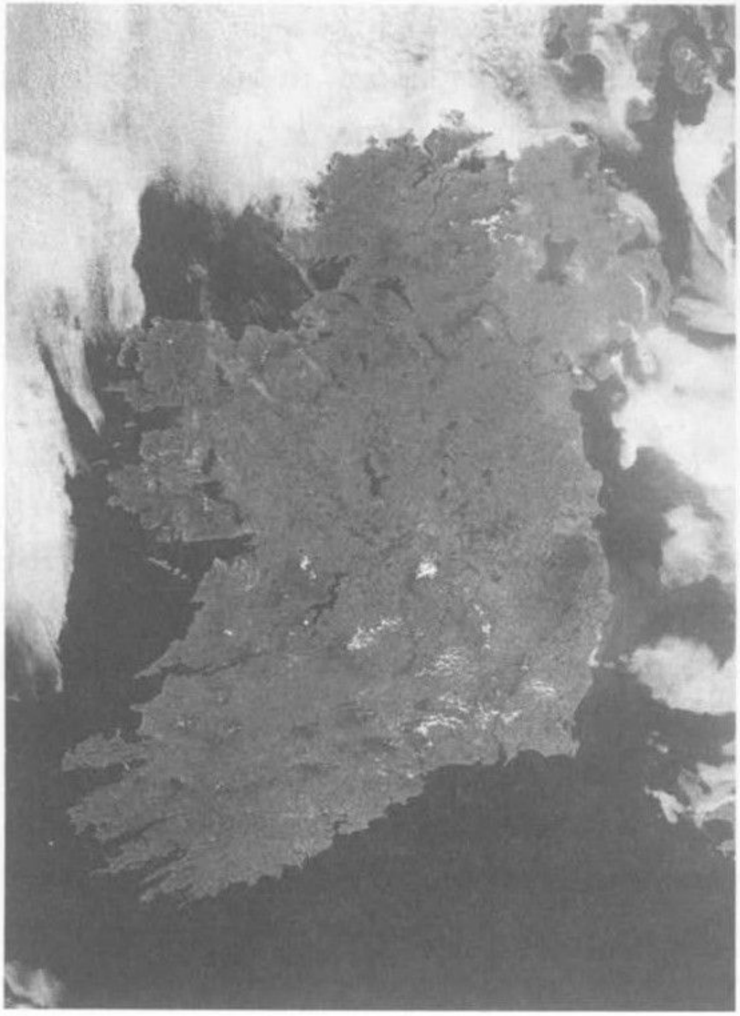
同前面描述的度量一样，度量统计复杂性也不容易，除非面对的系统可以解读为信息源。不过克鲁奇菲尔德、杨和他们的同事实际测量了一系列真实世界现象的统计复杂性，比如复杂晶体的原子结构[99]和神经元的激发模式[100]。

用分形维度量复杂性

前面讨论的复杂性度量都是基于信息论和计算理论的概念。但这并不是复杂性度量的唯一可能来源。还有人提出用动力系统理论的概念度量事物复杂性的方法。其中一个是用事物的分形维（fractal dimension）。要解释什么是分形维，还要先解释一下什么是分形。

分形最经典的例子是海岸线。从空中俯瞰下去，海岸线崎岖不平，有许多大大小小的海湾和半岛（图7.2，上图）。如果你下去沿着海岸线游览，它似乎还是一样的崎岖不平，只是尺度更小（图7.2，下图）。如果你站在沙滩上，或是以蜗牛的视角近距离观察岩石，相似的景象还是会一次又一次出现。海岸线在不同尺度上的相似性就是所谓的“自相似性”。

分形一词是由法国数学家曼德布罗特（Benoit Mandelbrot）提出来的，曼德布罗特认识到自然界到处都有分形——现实世界中许多事物都有自相似结构。海岸线、山脉、雪花和树是很典型的例子。曼德布罗特甚至提出宇宙也是分形的，[101]因为就其分布来说，有星系、星系团、星系团的聚团，等等。图7.3展示了自然界中一些自相似性的例子。



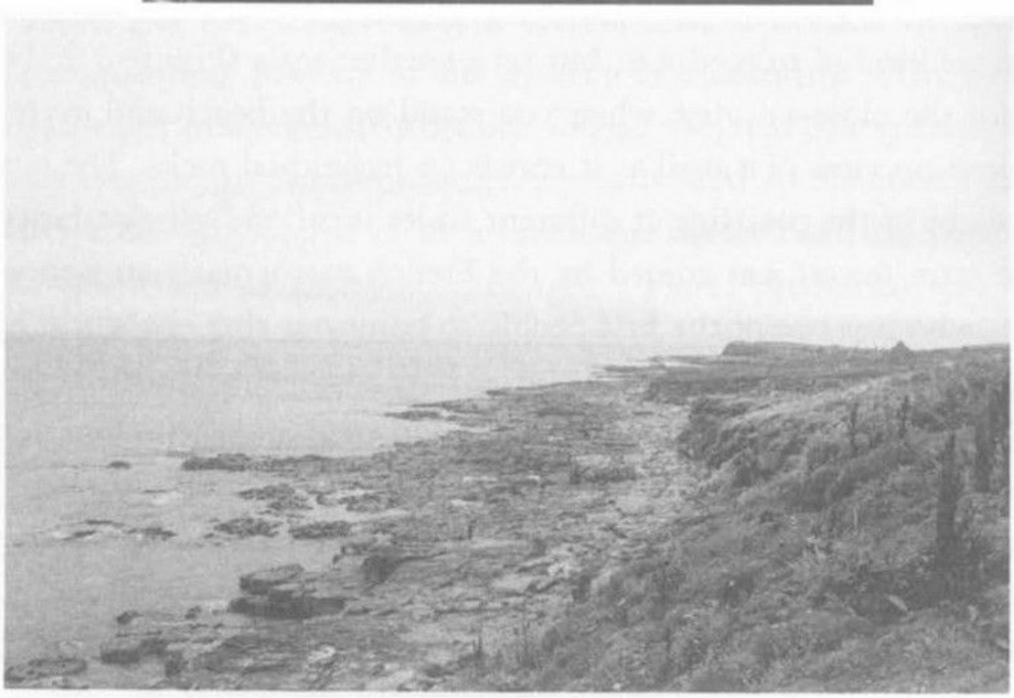


图7.2　上图：爱尔兰岛鸟瞰图片，其海岸线有自相似（分形）特征。下图：爱尔兰海岸线的局部图片。这个尺度上的崎岖结构与更大尺度上的崎岖结构类似[上图来自NASA可视地球（http：//visibleearth.nasa.gov/）。下图由Andreas Borchet拍摄，经Creative Commons许可使用（http：//creativecommons.org/licenses/by/3.0/）]

虽然人们对分形一词的意义有时候有不同理解，但一般来说分形指的是“在任何尺度上都有微细结构”的几何形状。[102] 许多让人感兴趣的分形具有自相似特性，海岸线就是这样的例子。第2章中逻辑斯蒂映射的分叉图（图2.10）也具有一定程度的自相似性。事实上，许多系统的混沌域（在逻辑斯蒂映射中是R大于3.57的部分）常被称为分形吸引子。

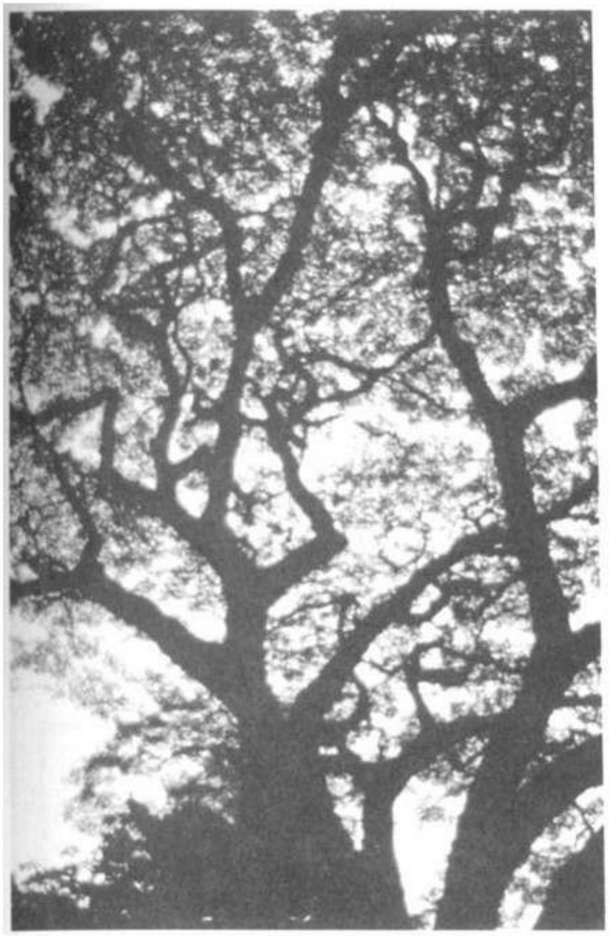
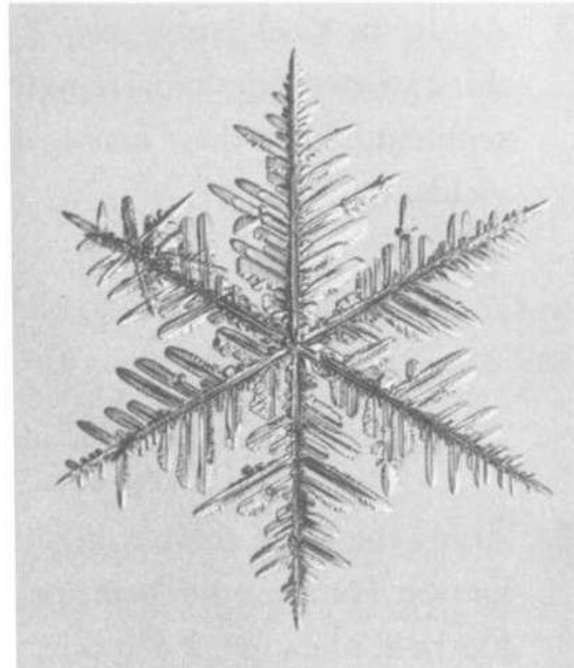
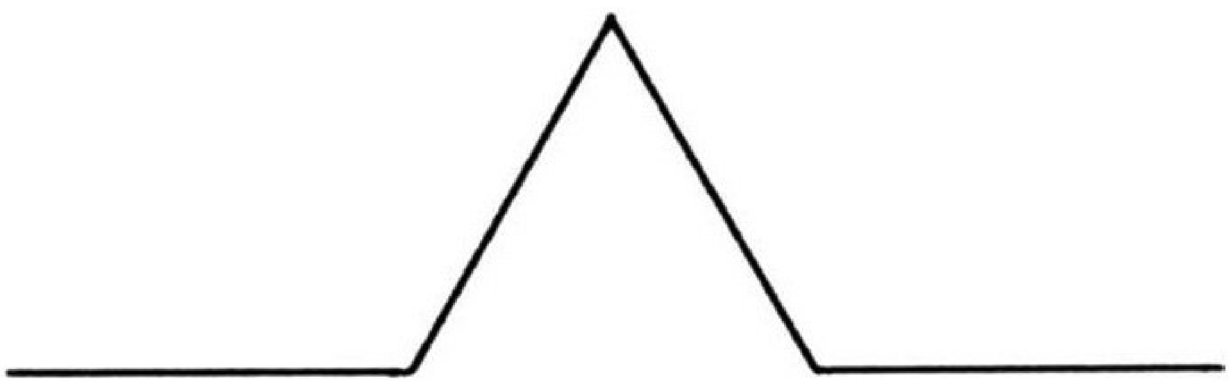




图7.3　自然界中一些分形结构的例子：树、雪花（显微镜放大）、星系团[树的照片来自美国国家海洋和大气管理局照片图书馆（National Oceanic and Atmospheric Administration Photo Library）。雪花照片来自http：//www.Snow Crystals.com，蒙Kenneth Libbrecht允许使用。星系团照片来自NASA太空望远镜科学研究所（NASA Space Telescope Science Institute）]

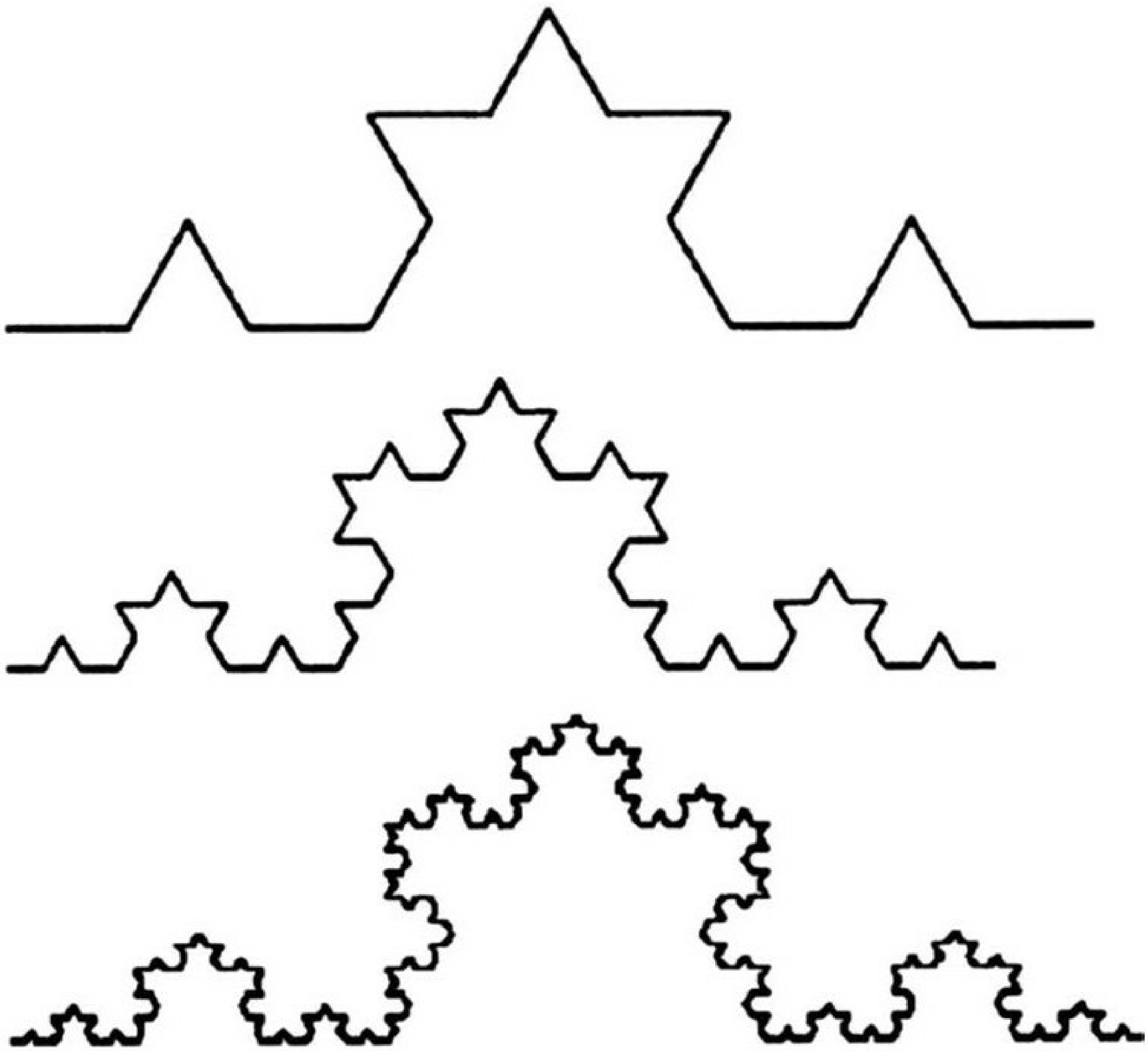
曼德布罗特等数学家为自然界中的分形设计了各种数学模型。其中一个很有名的例子是科赫曲线（Koch Curve，以发现这种分形的瑞典数学家命名）。科赫曲线是通过不断应用一条规则得出：

1.从一条直线段开始。

2.应用科赫曲线规则：“将每段线段等分成三段，中间一段替换为一个三角形的两条边，每一边都等于原线段的1/3。”因为只有一条线段，应用这个规则后变为：



3.对生成的图形再次应用科赫曲线规则，不断继续。下面是迭代了两次、三次和四次之后的情形：

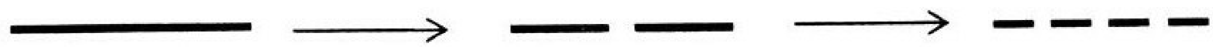


最后一张图有点像一条理想化的海岸线。（如果你向左旋转90度，然后斜着看，还真有点像阿拉斯加西海岸。）不过它是严格自相似的：曲线的部分，以及部分的部分，都与曲线整体是一样的形状。如果我们将科赫曲线规则应用无数次，图形在无数尺度上都将是自相似的——完美的分形。而真正的海岸线并不严格自相似。如果你观察海岸线的一小段，它并不与整段海岸线的形状完全一样，而是在许多方面相似（例如，蜿蜒崎岖）。另外，在真实世界中，自相似在无穷小的尺度上并不成立。为了简单起见，海岸线这类真实世界的结构通常被称为“分形”，但更严格的叫法应该是“类分形（fractal-like）”，特别是有数学家在场的时候。

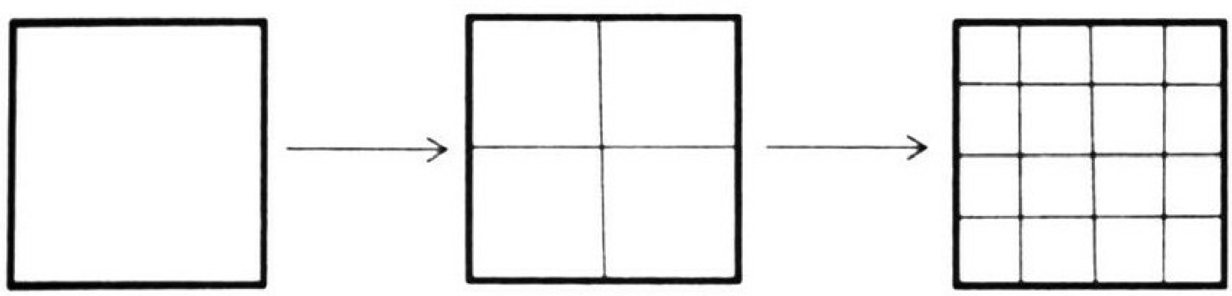
我们熟悉的空间维度的概念对于分形完全不适用。直线是1维，平面是2维，立方体是3维。那科赫曲线是几维呢？

首先我们来看看直线、正方形和立方体这些常规几何对象的维数到底指的是什么。

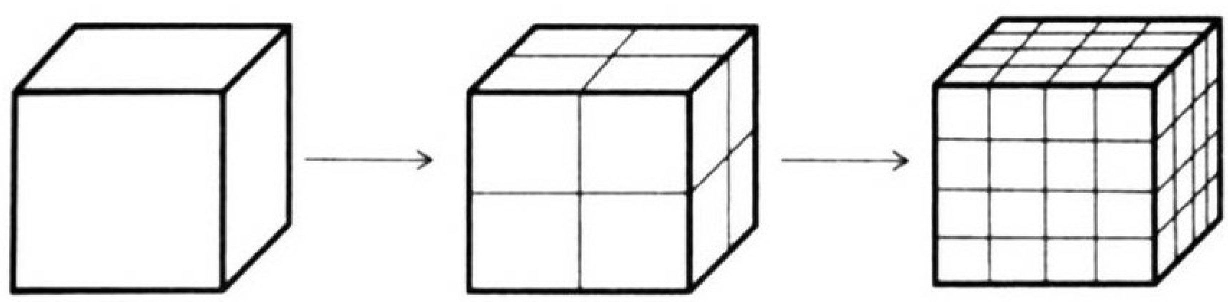
先来看看直线段。将其一分为二。然后将得到的线段再二分，每次都将各段线段一分为二：



每一次得到的图形都是由两个上次缩小一半的拷贝组成。再来看看正方形。从各边将其二分。然后将得到的正方形继续从各边二分，这样不断二分下去。



每次得到的图形都是由上次四分之一大小的4个拷贝组成。你可能已经猜到下面做什么了，将立方体从各边二分。将得到的立方体不断二分：



每次得到的都是由上次八分之一大小的8个拷贝组成。

这里已经能够看出维度的意义。一般而言，每次得到的图形都是由上次缩小的拷贝组成，而拷贝的数量则是2的维数次幂（$2^{\text{维数}}$）。对于直线，是2^1 =2个拷贝；对于正方形，是2^2 &=4个拷贝；对于立方体是2^3 =8个拷贝。类似的，如果不是二分，而是将各边三分，则每次得到的图形是上次的$3^{\text{维数}}$ 个拷贝。由此可以总结出一个规律：

> 将几何结构从各边分成x等份，不断重复这个过程。每次得到的将是前一次的$x^{\text{维数}}$ 个拷贝。

根据维数的这种定义，直线是1维，正方形是2维，立方体是3维。都没有问题。

现在将这个定义类推到科赫曲线。每次直线段都是之前的1/3长，而得到的则是之前的4个拷贝。根据前面的定义，应该是$3^{\text{维数}}$ =4。维数是多少呢？这里我们直接给出结果[103]（计算过程在注释中给出），根据前面的规律，维数约为1.26。也就是说，科赫曲线既不是1维也不是2维，而是介于两者之间。太奇怪了，分形的维数居然不是整数。这正是分形的奇特之处。

简而言之，分形维数[104]决定了物体的自相似拷贝的数量。同样，分形维也决定了随着层次的变化，物体总的大小（或者面积、体积）会如何改变。例如，如果你在每次应用规则后测量科赫曲线的总长度，你会发现每次长度增加为原来的4/3。只有完美的分形——可以缩小直至无穷——才有精确的分形维数。像海岸线这类真实世界的有穷类分形事物，我们只能测量近似的分形维数。

为了解释分形维数的直观意义，有过很多尝试。譬如，认为分形维表示了物体的“粗糙度”“凸凹度”“不平整度”或“繁杂度”；物体的“破碎”度；还有物体的“结构致密”程度。例如，比较爱尔兰（图7.2）和南非（图7.4）的海岸线，前者的分形维数比后者更高。

还有一种说法挺有诗意的，我很喜欢，即认为分形维数“量化了物体细节的瀑流”。[105] 也就是说，当你沿着自相似的瀑流越走越深时，它决定了你能看到多少细节。如果结构不是分形的，譬如平滑的大理石，你将它的结构不断放大，将不会出现有意思的细节。而分形则在所有层面上都有有趣的细节，分形维数一定程度上量化了细节的有趣程度

与你观察的放大率之间的关系。

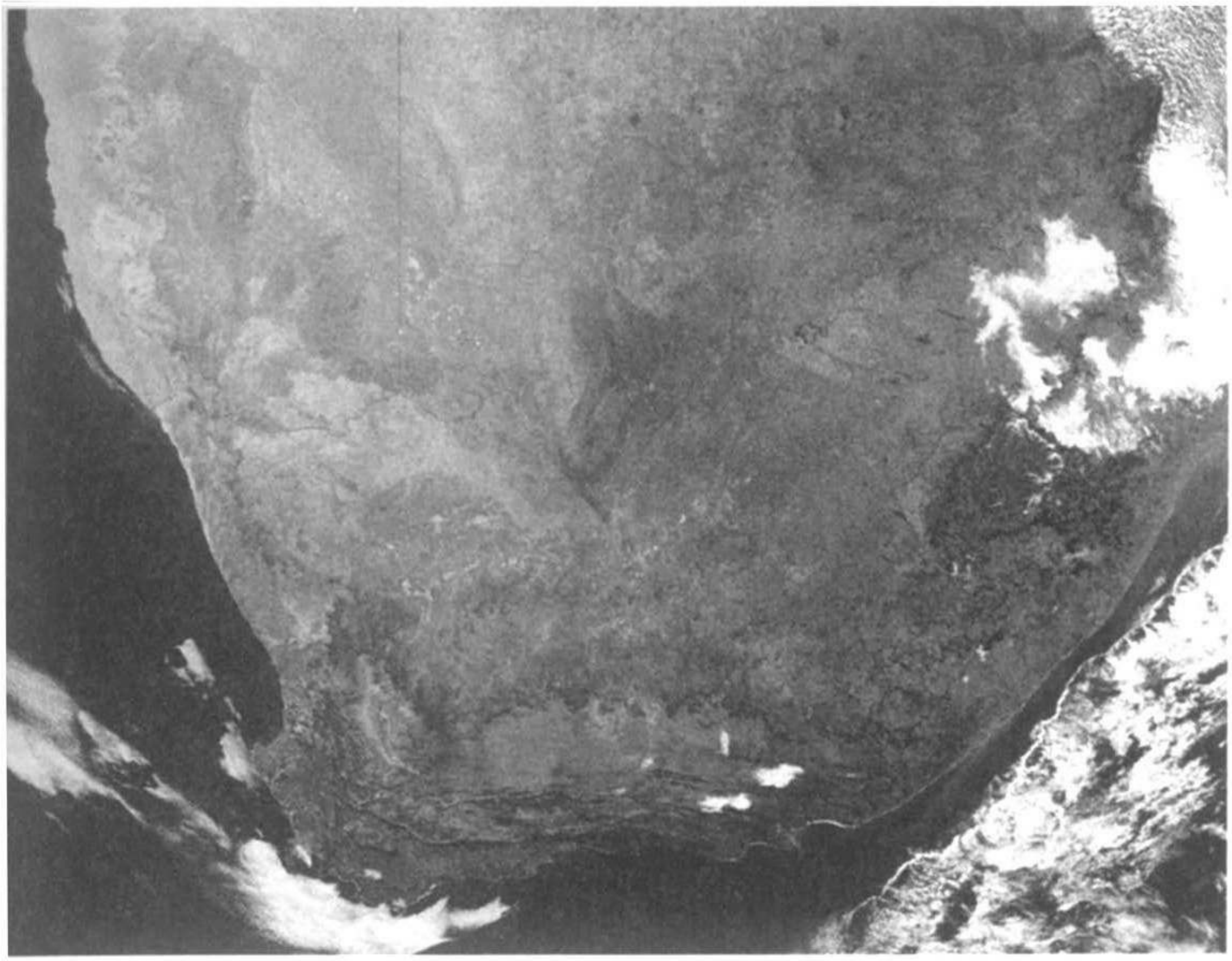


图7.4　南非海岸线[照片来自NASA可视地球（http：//visibleearth.nasa.gov）]

这也就是为何人们对用分形维数度量复杂性感兴趣，许多科学家都用其来度量真实世界的现象。不过，除了崎岖度和细节瀑流，还有许多其他种类的复杂性我们也希望能进行度量。

用层次性度量复杂性

1962年，西蒙（Herbert Simon）发表了一篇著名的文章——《复杂性的结构》[106]。文中西蒙提出一个系统的复杂性可以用层次度（degree of hierarchy）来刻画：“复杂系统由子系统组成，[107]子系统下面又有子系统，不断往下。”西蒙是位杰出的学者，他博学多识，既是政治学家、经济学家，又是心理学家，他的成就用一章的篇幅来讨论也不为过。

西蒙认为，复杂系统最重要的共性就是层次性和不可分解性。西蒙

列举了一系列层次结构的复杂系统——例如，身体由器官组成，器官又是由细胞组成，细胞中又含有细胞子系统，等等。某种程度上，这个观念与分形在所有尺度上都自相似类似。

不可分解性指的是，在层次性复杂系统中，子系统内部的紧密相互作用比子系统之间要多得多。例如，细胞内部的新陈代谢网络就比细胞之间的作用要复杂得多。

西蒙还认为，进化之所以能设计出自然界中的复杂系统，正是因为它们能像砖块一样被结合到一起——也就是说，具有层次性和不可分解性。细胞能够进化，从而成为高一级器官的建筑模块，组成的器官又可作为更高一级器官的建筑模块。西蒙认为复杂系统研究需要有一个“层次理论”。

许多人探讨了用层次性度量复杂性的可能途径。例如进化生物学家麦克西（Daniel Mc Shea）就一直想厘清生物随着进化复杂性增加的意义，他提出了一种层次标度，[108] 可以用来度量生物的层次度。麦克西的标度是用嵌套层次定义：高一级的对象嵌有低一级的对象作为组分。麦克西提出了以下嵌套的生物学层次：

层次1：原核细胞（最简单的细胞，例如细菌）；

层次2：层次1生物的聚合，例如真核细胞（更复杂的细胞，由原核细胞合并进化而来）；

层次3：层次2生物的聚合，例如所有多细胞生物；

层次4：层次3生物的聚合，例如昆虫群落和僧帽水母这样的群体生物。

随着嵌套继续，每一层可以说都比上一层更复杂。不过，就像麦克西所指出的，嵌套仅仅描述了生物的结构，而不涉及其功能。

麦克西用化石和现代生物的数据揭示了生物的最高层次随着进化而不断增加。因此可以用此说明复杂性随着进化不断增加，虽然在度量具体生物的层次度时对于何为“组成”或“层次”存在一些主观性。

还有很多度量复杂性的方法，在这里无法一一赘述。各种度量都抓住了复杂性思想的一些方面，但都存在理论和实践上的局限性，还远不能有效刻画实际系统的复杂性。度量的多样性也表明复杂性思想具有许多维度，也许无法通过单一的度量尺度来刻画。

2　计算机中的生命和进化

大自然逐渐将无生命的东西[109] 变成有生命的动物，其间的界线无法分辨。

——亚里士多德，《动物史》（History of Animals ）

我们都本能地知道生命是什么[110] ：它是可以用来吃的，可以爱的，甚至可能是致命的。

——洛夫洛克（James Lovelock），

《盖亚时代》（The Ages of Gaia ）

第8章　自我复制的计算机程序[111]

生命是什么

第5章介绍了一些生命进化的思想史。但是有一些问题没有提到，例如生命是如何起源的？生命的要素到底是什么？这些都是科学界最具争议的问题，至今没有确定的答案。在这里我不讨论前一个问题，不过复杂系统对此有一些让人着迷的研究。[112]

生命到底是什么，这是一个经久不衰的问题。不管是大众还是科学家，对此都没有达成共识。像“生命是如何诞生的”或者“生命在其他星球上是什么样子”总是能激起热烈的讨论，甚至引起敌意。

创造人工生命的想法也由来已久，至少可以追溯到两千年前的石人传说（Golem）和奥维德（Ovid）的皮格马利翁，[113] 19世纪又有福兰克斯坦的怪兽（Frankenstein’s monster）的故事，更不要说现在的《刀锋战士》和《黑客帝国》这些影片，以及《模拟生命》（Sim Life）这类计算机游戏。

科幻作品提出了一个新版本的“生命是什么”的问题：计算机和机器人可以被认为有生命吗？这个问题将计算、生命和进化的问题联系到一起。

如果你问10个生物学家什么是生命的10个要素，每次得到的答案都会不一样。可能大部分会包括自主、新陈代谢、自我复制、生存本能，还有进化和适应。我们能不能将这些过程机械化，并用计算机来实现呢？

许多人认为绝对不可能，理由如下：

自主：计算机本身什么都做不了；只能执行程序的指令。

新陈代谢：计算机无法像生物那样从环境中获取能量；它们必须由人提供能源（例如电力）。

自我复制：计算机不能复制自身；要复制自身就必须包含对自身的描述；而这个描述又包含其本身的描述，这样反复无穷。

生存本能：计算机不关心自己能不能生存，它们也不关心自己是不是成功。（有一次听一位杰出的心理学家的讲座，他在谈论计算机象棋程序时说：“就算深蓝赢了卡斯帕罗夫，它也不会有快乐的感觉。”）

进化和适应：计算机本身无法进化或适应；它只能严格依照程序员预先设定的方式变化。

虽然还有很多人相信这些观点，但它们都在人工生命领域中[114]以各种方式被否定了。人工生命关注的是在计算机中仿真或“创造”生命。在这一章和下一章我会讨论这些，它们与达尔文主义的自我复制和进化密切相关。

计算机中的自我复制

自我复制的观点非常数学化：它认为计算机中的自我复制会导致无穷反复。

我们先来看看计算机自我复制问题最简单的形式：写一段程序打印

其自身。

在后面我们用一种简单的计算机语言，这样不用是程序员也能看懂。（实际上是一种伪代码，有一些真正的计算机语言所没有的命令，但是是合理的，只是让事情更简单一些。）

下面来试一下，首先从程序的名字开始：

program copy

然后加一条指令将程序的名字打印出来：

program copy

print（“program copy”）

指令print只是简单地将引号之间的字符显示到屏幕上然后换行。现在增加一条指令将第二行打印出来：

program copy

print（“program copy”）

print（“print（“program copy”）”）

因为要输出程序自身的完整复制，第二个print指令的引号中带了第一个print指令前的四个缩进符和后面的一对引号（print指令输出最外侧的引号之间的所有字符，包括引号）。现在又要加一条指令来输出第三行：

program copy

print（“program copy”）

print（“print（“program copy”）”）

print（“print（“print（“program copy”）”）”）

现在你可能已经看出来，这种策略——每条指令输出上一条指令的拷贝——是如何导致无穷反复的。怎样才能避免这种情况呢？在继续往下读之前，你可以花点时间自己试一试能不能解决这个问题。

这个看似简单的问题其实关系到第4章介绍过的哥德尔和图灵的工作。解决方法同时也包含了生物系统本身绕开无穷反复的基本途径。20世纪匈牙利数学家冯·诺依曼在研究一个更复杂的问题时，首先发现了这个问题的答案。

冯·诺依曼是量子力学、经济学等多个领域的先驱，也是最早设计电子计算机的人之一。他的设计中包含中央处理单元和可以存储程序和数据的随机存取存储器。这些至今仍然是现代计算机的基础。冯·诺依曼也是最早深刻认识到计算和生物之间联系的科学家之一。他在生命最后的岁月里一直致力于解决机器如何才能复制自身的问题。他给出了第一个能自我复制的机器的完整设计。我在后面展示的自复制计算机程序就是受他的“自复制自动机”启发，并以简化的方式阐释其基本原则。

在介绍自复制程序之前，还要先解释一些会用到的编程语言的相关知识。

参考图8.1给出的计算机存储器示意图。在我们高度简化的例子中，计算机存储器由有编号的位置或“地址”组成，图中编号为1—5，依次往后。各位置中有一些字符。这些字符可以作为程序的指令或程序使用的数据。如果执行当前存储的程序，会显示输出：

Hello, world！

Goodbye.

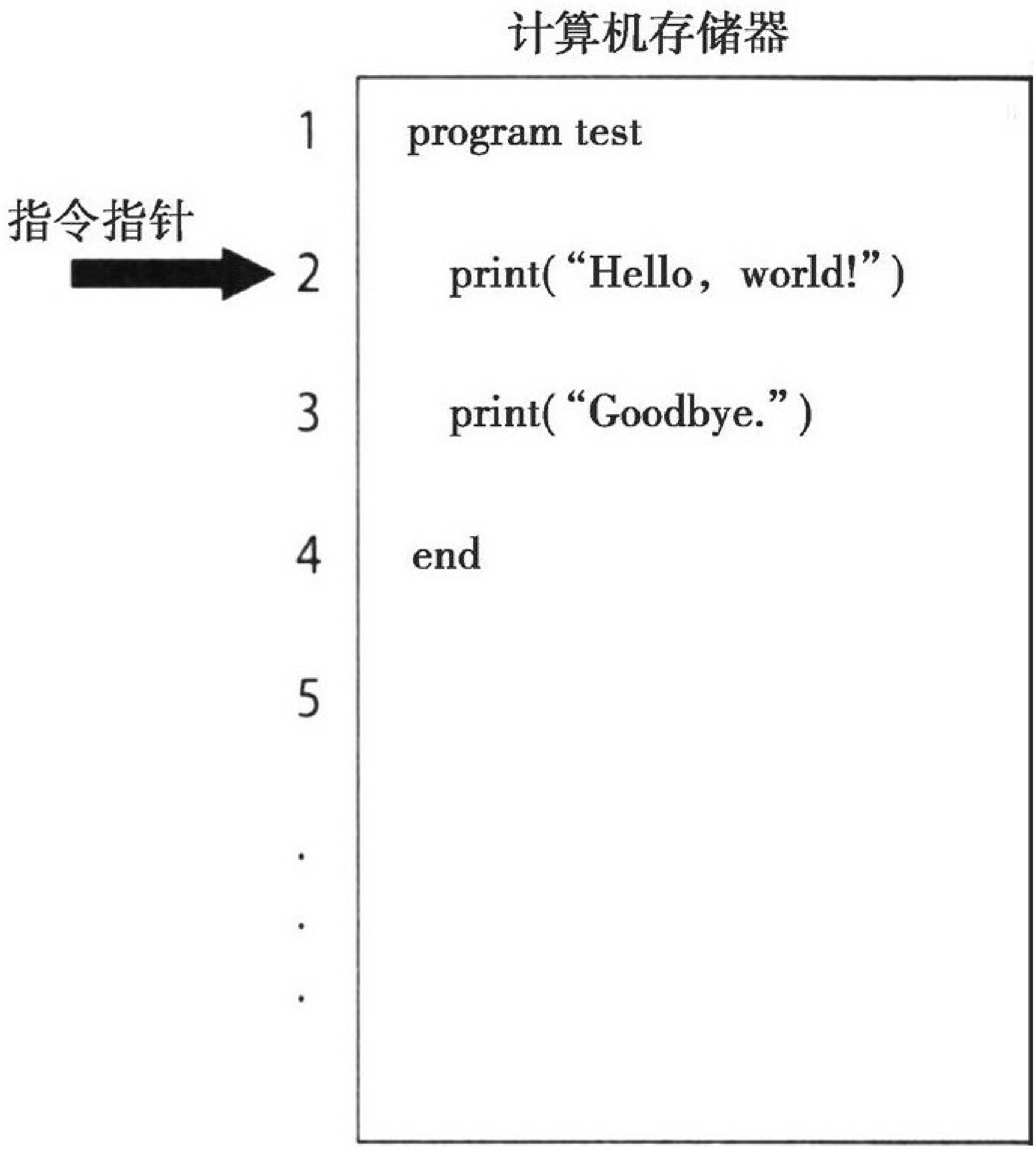


图8.1　计算机存储器简化示意图，位置用1—5依次编号，其中4个存储有程序。指令指针指向计算机当前存储的指令。有些指令行前面带有空格，在执行的时候会被忽略

要执行程序，计算机要有一个“指令指针”——同样存储在存储器中的一个数字，记录当前执行的指令在存储器中的位置。指令指针——简记为ip——最初设为程序第一行的存储地址。我们称之为“指向”那条指令。在计算的每一步ip指向的指令会被执行，ip加1。

例如，在图8.1中，ip的值为2，也就是说指向的是print（“Hello, world！”）。

我们称ip为变量，因它的值随着计算的进行而不断变化。

还可以定义变量line[n]表示地址n中的字符串。例如，指令print（line[2]）会显示输出：

print（“Hello, world！”）

此外，我们的编程语言中还包括loop指令（循环）。例如，下面的程序代码：

```
x=0

loop until x=4

{

print（“Hello, world！”）

x=x+1

}
```

会输出：

Hello, world！

Hello, world！

Hello, world！

Hello, world！

大括号之间的代码会反复执行直到循环结束条件（这里是x=4）满足。变量用作计数器——从0开始，每循环一次加1。增加到4时循环停止。

现在可以来看看自我复制程序了，程序完整展现在图8.2中。理解一段程序最好的办法就是手工推演，也就是一行一行跟踪程序的运行。

假设图8.2中的程序被加载到内存中，然后假设有人在计算机命令提示符后键入selfcopy，计算机就会开始执行程序selfcopy。译码器——操作系统的一部分——会将指令指针设为1，指向程序名。然后ip会下移，逐行执行各条指令。

在地址2处变量L被设为ip-1。而ip是当前执行指令的位置。因此当执行第2行时，ip设为2，L设为2-1=1。（注意虽然随着指令执行会不断变化，但L在重置之前会一直等于1直到其被重置。）接着会进入循环，直到line[L]等于字符串end。前面说了line[L]等于内存中地址L处的字符串。目前L等于1，line[L]等于字符串programselfcopy，不等于字符串end，因此循环不会停止。在循环中，会显示输出line[L]，并将L加1。最初，L=1，显示输出program selfcopy；然后L被置为2。

计算机内存

```
program selfcopy
    L=ip−1
    loop until line[L]= “end”
    {
        print（line[L]）
        L=L+1
    }
    print（ “end” ）
end
```

图8.2　自我复制的程序

现在，line[L]指的是程序第二行，即L=ip-1，仍然不等于end，因此循环会继续。这样程序就会被逐行输出。尤其有意思的是第5行：当L=5时，执行第5行，指令print（line[L]）会显示输出其自身。当L=9时，line[L]等于end，循环终止。这时已经显示输出了1—8行。指令指针指向第8行（紧跟在循环后面的指令），执行时显示输出字符串“end”结束自我复制。在这个程序中自我复制的本质，是用两种方式来使用内存中的信息：既作为执行的指令，又作为这些指令使用的数据。正是对信息的双重使用让我们得以避开前面尝试自我复制程序时遇到的那种无穷反复。

自我复制程序的深层意义

信息的双重使用是哥德尔悖论的核心，他的自指句子“这个命题是不可证的”体现的正是这一点。

理解这个需要要点把戏。首先，请注意这个句子同其他句子一样，可以从两个角度来看：①视为句子中包括的文字、空格符和标点组成的字符串；②视为字符串所代表的意义，同语言使用者的解读一样。

为了明确起见，我们将句子的字符串本身记为S。也就是说，S=“这个命题是不可证的”。现在可以陈述的一些命题：例如它包含9个字，1个句号。

句子的意义则记为M。我们可以将M重写为“命题S是不可证的”。某种程度上，你可以将M视为“指令”，而将S视为指令操作的数据。怪异（而又神奇）的是，数据S与指令M是同一个东西。哥德尔之所以能将句子转化为数学中的悖论，一个主要原因就是他能将M表示成数学命题，将S表示成编码那个数学命题字符串的数字。

这就是背后的把戏。侯世达在《哥德尔、艾舍尔、巴赫——集异璧之大成》一书中，对字符串与字符串的意义之间的区别，以及自指所导致的悖论，进行了详细而有趣的讨论。

与之类似，对信息的双重使用也是图灵对停机问题不可判定性证明的关键。还记得第4章的H和H′吗？记得H′是如何作用于自身的编码吗？同这里的自复制程序一样，H′也是以两种方式被使用：解释为程序，同

时也作为程序的输入。

DNA的自我复制

现在你可能在想，我们又回到了让人头痛的抽象逻辑王国。别急，我们马上就会回到现实世界。真正让人惊奇的是对信息的双重使用竟然也是DNA复制自身的关键。在第6章我们了解到，DNA是由核苷酸序列组成。特定的子序列（基因）编码构成蛋白质的氨基酸，其中包括解开双螺旋和用信使RNA、转运RNA、核糖体等复制各股DNA的酶（特定种类的蛋白质）。作一个宽泛的类比，对执行复制的酶进行编码的DNA序列大致上对应于自复制程序的代码。这些DNA中的“代码”在产生酶和作用于DNA自身时就相当于被执行，DNA本身则相当于被解开和复制的数据。

不过你可能已经注意到了我埋下的伏笔。在自复制程序和DNA的自复制之间有一个重要的差别。自复制程序需要有一个解释器来执行它：指令指针依次指向各行代码，然后由操作系统来执行它们（存取ip和L等中间变量，显示输出字符串，等等）。执行器完全外在于程序本身。

而在DNA的情形中，构建“解释器”——信使RNA、转运RNA，核糖体和所有用于蛋白质合成的成分——的指令也一起编码在DNA中。也就是说，DNA不仅包含自我复制的“程序”（例如用来解开和复制DNA的酶），同时也编码了它自己的解释器（将DNA转译成酶的细胞器）。

冯·诺依曼的自复制自动机

冯·诺依曼最初的自复制自动机（冯·诺依曼只给出了数学描述，并没有真的建造）也是既包含有自我复制的程序也包含解释自身程序的机制。因此是完整的自我复制机器。这也解释了为何冯·诺依曼的构想要比我的自我复制程序复杂得多。冯·诺依曼提出构想是在20世纪50年代，当时生物自我复制的机制还没有被完全理解，这更加表明了冯·诺依曼天才的洞察力。冯·诺依曼对自动机的设计以及其正确性的数学证明在他去世前已基本完成。1957年，他因癌症去世，年仅53岁，可能是因为在参与研制原子弹时受到了核辐射。冯·诺依曼的同事巴克斯（Arthur Burks）完成了最后的证明。1966年，巴克斯将全部成果编辑成为《自复制自动机理论》（*Theory of Self-Reproducing Automata*）一

书出版。[115]

冯·诺依曼设计的自复制自动机是人工生命科学真正的先驱之一，从原则上证明了自我复制的机器的确是可能的，并且提供了自我复制的“逻辑”，后来证明其与生物的自我复制机制惊人的相似。

冯·诺依曼认识到这个结论具有深远的影响。他担心公众对这种自复制机器的反应，不愿看到大众媒介报道“未来这种自复制机器的可能性”[116]。可惜好景不长。1999年，计算机科学家库兹韦尔（Ray Kurzweil）和莫拉韦茨（Hans Moravec）在《灵魂机器的时代》（*The Age of Spiritual Machines*）和《机器人》（Robot）这两本书中鼓吹了这种具有超级智能并且能自我复制的机器人的可能性，他们认为这种机器人在不远的将来就会被制造出来，他们的书并非虚构，[117]但是相当牵强。2000年，Sun公司的创始人之一乔伊（Bill Joy）在《连线》（Wired）杂志上发表了一篇后来很有名的文章[118]——《为何未来不需要我们》，文中描述了自复制纳米机器的可能性。目前这些预言都还没有应验。不过复杂的自复制机器也许很快就会成为现实：康奈尔大学的机器人专家利普森（Hod Lipson）和他的同事已经制造出了一些简单的自复制机器人[119]。

冯·诺依曼

冯·诺依曼（图8.3）是20世纪科学和数学领域最重要的人物，而且很有趣，在这里值得多说几句。如果你还不知道他的话，应该去了解一下。



图8.3　冯·诺依曼（1903—1957）（美国物理学会西格尔图像档案）

不管和谁比，冯·诺依曼都是真正的天才。在相对短暂的一生中，他至少在6个领域作出了基础性的贡献：数学、物理、计算机科学、经济学、生物学和神经科学。人们说起他的故事时，总是忍不住摇摇头，惊叹如此的天才是不是真的是人类能做到的。我很喜欢他的故事，希望在这里与你们分享。

冯·诺依曼出生在匈牙利，小名琼尼。与爱因斯坦和达尔文的大器晚成不同，冯·诺依曼从小就是神童。据说他六岁时就能心算八位数除法。（很久他才发现不是所有人都能做到这一点；他的一本传记中有这样一个故事："6岁时，有一次他母亲在他面前显得心不在焉，[120] 他问妈妈：'你在算什么？'"）当时他还能和父亲谈论古希腊。

冯·诺依曼18岁进入大学，开始是在布达佩斯，后来又去了德国和瑞士。最初他选择的是化学工程这样的实用课程，但还是无法离开数学。23岁时，因为在数理逻辑和量子力学做出的基础性工作，他获得了数学博士学位。他的工作做得太漂亮了，5年后他就获得了世界上最好的学术职位——加入新成立的普林斯顿高等研究院（IAS），爱因斯坦和哥德尔也是这里的成员。

研究院没有看走眼。此后10年，冯·诺依曼开创了博弈论的研究（写出了被称为"有史以来最好的数理经济学论文" [121] ），设计了第一台可编程计算机的原理架构（EDVAC，他为这台计算机写的报告被称为"计算和计算机领域有史以来最重要的文献" [122] ），还在第一颗原子弹和氢弹的研制中做出了重要贡献。此后他又致力于研究自复制自动机以及计算机逻辑与大脑运作机制之间的关系。冯·诺依曼在政治上也很活跃（他的立场十分保守，持强烈的反共产主义观点），后来还成为原子能委员会的成员，这个委员会为美国总统在核武器政策方面提供咨询。

除了冯·诺依曼，匈牙利还有一批年龄相仿的科学家后来都成了举世闻名的学者，这被称为"匈牙利现象"。这个群体中包括西拉德，第3章我们已经见过他，物理学家维格纳（Eugene Wigner）、特勒（Edward Teller）和伽柏，数学家厄多斯（Paul Erd s）、科蒙尼（John Kemeny）和拉克斯（Peter Lax）。许多人都奇怪为何当时会聚集这么

多耀眼的天才。据冯·诺依曼的传记作者麦克雷（Norman Mac Rae）说："匈牙利6位诺贝尔奖获得者有5位[123]是生于1875年到1905年间的犹太人，有一次诺贝尔奖得主维格纳被问道，为何在他那一代匈牙利涌现了这么多天才，结果他回答说，他不明白这个问题，匈牙利当时只出现过一位天才，那就是冯·诺依曼。"

冯·诺依曼在许多方面都领先于他的时代。他的目标同图灵类似，想发展信息处理的一般理论，既包括生物也包括技术。他在自复制自动机上的工作就是这项计划的一部分。冯·诺依曼与控制论社团也联系紧密——这是一个由科学家和工程师组成的交叉学科研究群体，致力于研究各种自然界和人工复杂适应系统的共性。现在的"复杂系统"研究的前身就是控制论和系统科学。在最后一章我将进一步探讨它们之间的联系。

冯·诺依曼对计算的兴趣在高等研究院并不是一直都受欢迎。在结束关于EDVAC的工作之后，冯·诺依曼带领几位计算机专家在IAS设计和改进EDVAC的后续机型。这个系统被称为"IAS计算机"；它的设计是后来IBM早期计算机的原型。然而IAS一些纯理论科学家和数学家对此感到不快，认为在这个纯洁的象牙塔中不应进行这种实用性研究，对于冯·诺依曼领导一组气象学家在IAS用这台计算机开展的第一项应用天气预报可能更加反感。一些纯粹主义者认为这类研究不适合研究院。就像IAS的物理学家戴森（Freeman Dyson）说的，"（IAS）数学院[124]分为三个群体，一群是纯数学，一群是理论物理，冯·诺依曼教授单独是一群"。冯·诺依曼去世后，IAS的计算机项目被终止了，IAS的研究人员进行了一场"让研究院没有任何实验科学[125]，没有任何实验室"的运动。戴森称之为"自大狂的报复"[126]。

第9章　遗传算法

在对"机器能否复制自身"的问题给予肯定回答后，冯·诺依曼很自然地想让计算机（或计算机程序）复制自己和产生变异，并在某种环境中为生存竞争资源。这就会遇到前面提到的"生存本能"以及"进化和适应"的问题。可惜的是冯·诺依曼还没有研究进化问题就去世了。

其他人很快就开始继续他留下的工作。20世纪60年代初，一些研究

团体开始在计算机中进行进化实验。这些研究现在统称为进化计算[127]（evolutionary computation）。其中最为著名的是密歇根大学的霍兰德和他的同事、学生进行的遗传算法（genetic algorithms）研究。

霍兰德（图9.1）可以说是冯·诺依曼的学术徒孙。霍兰德的博士导师是哲学家、逻辑学家和计算机工程师巴克斯，巴克斯曾协助冯·诺依曼研制EDVAC，并且完成了冯·诺依曼没有完成的自复制自动机研究。在结束EDVAC的工作之后，巴克斯在密歇根大学获得了哲学教职，并成立了计算机逻辑小组，这是一个由对计算机基础以及广义信息处理感兴趣的教师和学生组成的松散团体。霍兰德到密歇根大学攻读博士学位，开始是学数学，后来转到新成立的“通信科学”系（后来改称“计算机与通信科学”），这可能是世界上第一个真正的计算机科学系。几年后，霍兰德成为系里第一个博士学位获得者，他也是世界上第一个计算机科学博士。很快他就留校成了计算机系的教授。

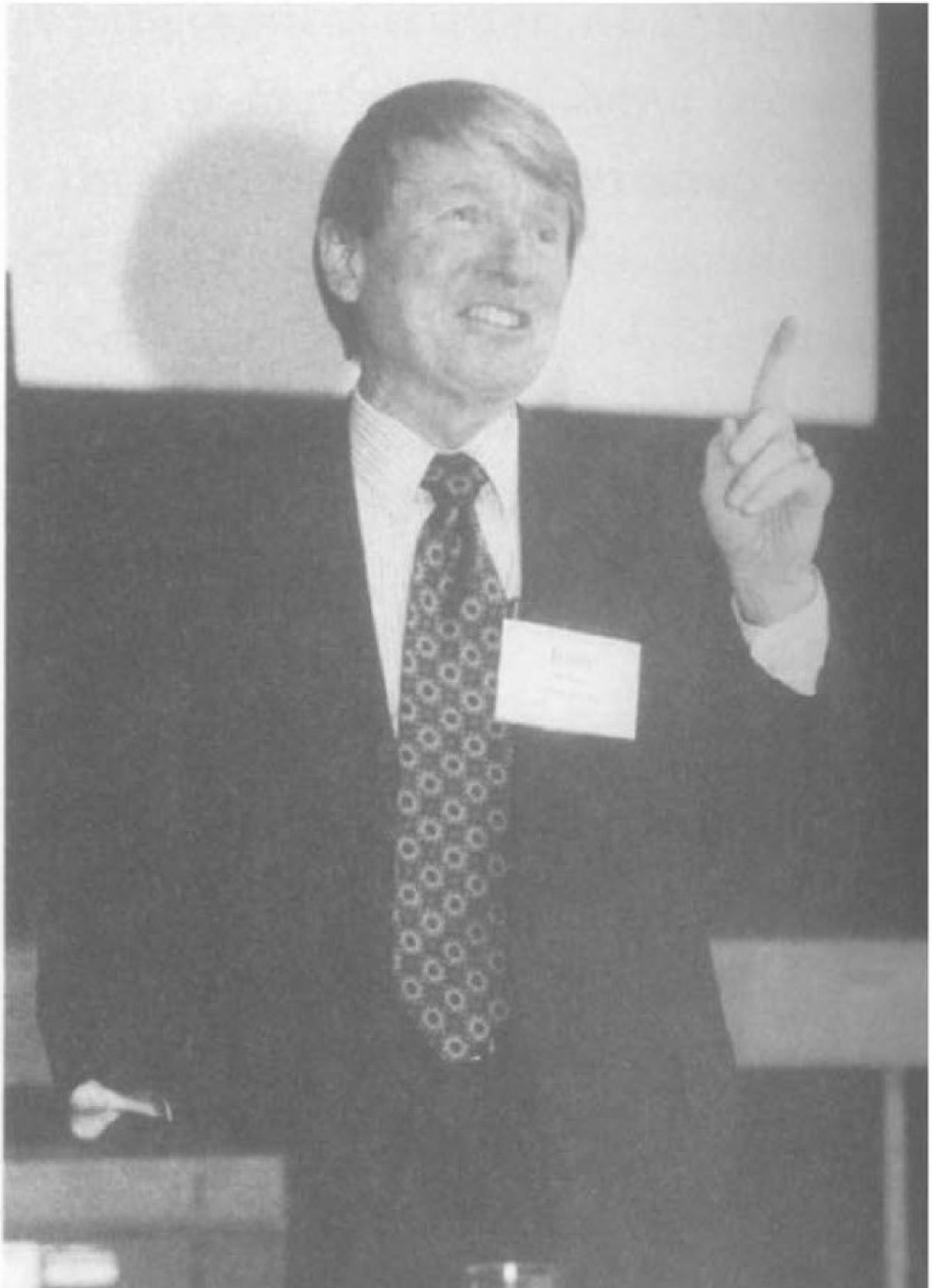


图9.1　霍兰德（圣塔菲研究所版权所有，经许可重印）

霍兰德在读费希尔（Ronald Fisher）的名著《自然选择的遗传理论》（*The Genetical Theory of Natural Selection*）时被达尔文的进化论深深吸引。同费希尔（和达尔文）一样，进化与农产养殖之间的相似也给霍兰德留下了深刻印象。但他是从计算机科学的角度来思考这种相似性："这就是遗传算法的由来。[128] 我想到，是不是可以像繁育良种马和良种玉米那样繁殖程序。"

霍兰德的主要兴趣在于适应现象——生物如何进化以应对其他生物和环境变化，计算机系统是不是也可以用类似的规则产生适应性。他在1975年的著作《自然和人工系统的适应》（*Adaptation in Natural and Artifcial Systems*）中列出了一组适应性的普遍原则，并且提出了遗传算法的构想。

我第一次了解遗传算法是在密歇根大学研究生院，当时我选了霍兰德基于他的书开的一门课。我马上就被"进化的"计算机程序的思想吸引住了。（同赫胥黎一样，我的反应是："我怎么没想到，真是太蠢了！"）

遗传算法菜谱

算法其实就是图灵说的明确程序，就好比做菜的菜谱：一步一步将输入变成输出。

对于遗传算法（GA），期望的输出就是特定问题的解。比如，你需要编写一个程序控制机器人清洁工在办公楼拾垃圾。你觉得编这个程序太费时间，就委托遗传算法替你将这个程序演化出来。因此，期望的GA输出就是能让机器人清洁工完成拾垃圾任务的控制程序。

GA的输入包括两部分：候选程序群体和适应性函数。适应性函数用来确定候选程序的适应度，度量程序完成指定任务的能力。

候选程序可以表示成位、数字或符号组成的字符串。后面我会给出一个机器人控制程序表示成数字字符串的例子。

在机器人清洁工的例子中，候选程序的适应度可以定义为机器人在

给定时间内清扫的面积，这是由程序决定的，越大越好。

下面是GA菜谱。

将下面的步骤重复数代：

1.生成候选方案的初始群体。生成初始群体最简单的办法就是随机生成大量“个体”，在这里个体是程序（字符串）。

2.计算当前群体中各个个体的适应度。

3.选择一定数量适应度最高的个体作为下一代的父母。

4.将选出的父母进行配对。用父母进行重组产生出后代，伴有一定的随机突变概率，后代加入形成新一代群体。选出的父母不断产生后代，直到新的群体数量达到上限（即与初始群体数量一样）。新的群体成为当前群体。

5.转到第2步。

遗传算法的应用

前面描述GA时似乎很简单，但是遗传算法已被用于解决科学和工程领域的许多难题，甚至应用到艺术、建筑和音乐。

应用之广泛从下面这些问题可见一斑：通用电气将GA用于飞行器的部分自动化设计，[129] 洛斯阿拉莫斯国家实验室用GA分析卫星图像，[130] 约翰·迪尔（John Deere）公司将GA用于自动化生产线的调度，[131] 德州仪器（Texas Instruments）则用GA来设计计算机芯片。[132] GA还在2003年的电影《指环王：王者归来》（*The Lord of the Rings*：*The Return of the King*）中被用于生成逼真的动画马匹，[133] 为电影《特洛依》（Troy）生成逼真的演员替身动画特效。[134] 许多制药公司用GA来辅助发现新药。[135] GA也被一些金融组织用于各种场合：识别交易欺诈[136]（伦敦股票交易所）、分析信用卡数据[137]（第一资本金融公司，Capital One）、预测金融市场[138] 和优化证券投资组合[139]（第一象限公司，First Quadrant）。20世纪90年代，互动式遗传算

法创造的艺术作品[140]在巴黎蓬皮杜中心（Georges Pompidou Center）等多个博物馆展出。这些只是遗传算法应用的一小部分例子。

进化的罗比，易拉罐清扫机器人

下面我们用一个更详细的简单例子[141]来进一步阐释GA的主要思想。我有一个叫“罗比”的机器人，它的世界（用计算机模拟，但是很脏乱）是二维的，到处是丢弃的易拉罐。我将用遗传算法为罗比进化出一个“脑”（即控制策略）。

罗比的工作是清理它的世界中的空易拉罐。罗比的世界由10×10的100个格子组成（图9.2）。罗比在位置（0，0）。我们可以假设周围围绕着一堵墙。许多格子中散落着易拉罐（不过每个格子中的易拉罐不会多于一个）。

罗比不是很聪明，看得也不远，他只能看到东南西北相邻的4个格子以及本身所在格子中的情况。格子可以是空的（没有罐子），或者有一个罐子，或者是墙。例如，在图9.2中，罗比位于格子（0，0），看到当前格子是空的，北面和西面是墙，南面的格子是空的，东面的格子中有一个罐子。

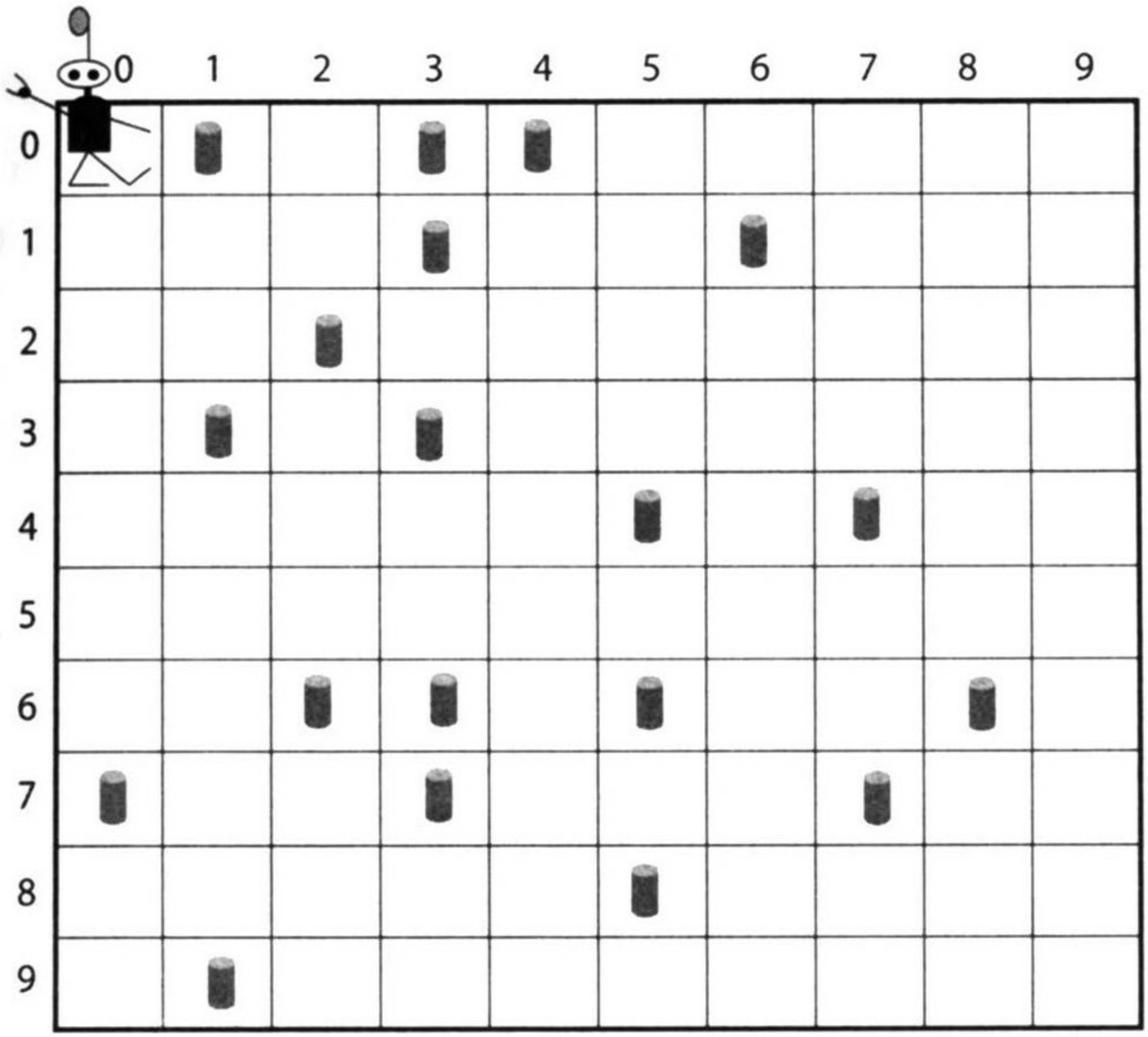


图9.2　罗比的世界。10×10的格子，散落着一些易拉罐

每次清扫工作罗比可以执行200个动作。动作可以是以下7种：往北移动、往南移动、往东移动、往西移动、随机移动、不动、收集罐子。每个动作都会受到奖赏或惩罚。如果罗比所在的格子中有罐子并且收集起来了，就会得到10分的奖赏。如果进行收集罐子的动作而格子中又没有罐子，就会被罚1分。如果撞到了墙，会被罚5分，并弹回原来的格子。

显然，罗比尽可能地多收集罐子，别撞墙，没罐子的时候别去捡，得到的分数就最高。

这个问题很简单，人工为罗比设计一个好策略可能也不是很难。不

过，有了遗传算法我们就可以什么也不用干，我们只需要等着计算机替我们进化出来。下面我们用遗传算法来为罗比进化出一个好策略。

第1步是搞清楚我们想要进化的到底是什么。也就是说，策略具体指的是什么？一般来说，策略指的是一组规则，规则给出了在各种情形下你应当采取的行动。对于罗比，它面对的“情形”就是它看到的：当前格子以及东南西北四个格子中的情况。对于“在各种情形下怎么做”的问题，罗比有7种可能选择：北移、南移、东移、西移、随机移动、不动、收集罐子。因此，罗比的策略可以写成它可能遇到的所有情形以及面对每种情形应当采取的行动。

有多少种可能的情形呢？罗比可以看到5个格子（当前格子、东、南、西、北），每个格子可以标为空、罐和墙。这样就有243种可能情形。[142] 其实还没有这么多，因为有许多不可能的情形，例如当前位置不可能是墙，也不可能四面都是墙，等等。不过，因为我们很懒，不想费劲找出所有不可能的情形，因此我们会列出所有243种情形，只要知道其中一些永远也不会遇到就行了。

表9-1是一个策略的例子——只是策略的局部，完整策略太长了，不方便列出来。

表9　1一个策略的例子

情形					行动
北	南	东	西	中	
空	空	空	空	空	往北移动
空	空	空	空	罐	往东移动
空	空	空	空	墙	随机移动
空	空	空	罐	空	清扫罐子
		…			
墙	空	罐	墙	空	往西移动
		…			
墙	墙	墙	墙	墙	不动

罗比在图9.2中的情形是：

北	南	东	西	中
墙	空	罐	墙	空

要知道下一步怎么做，罗比只需要查看策略表，查到对应的行动是往西移动。因此它往西移动一格，结果一头撞到墙上。

我没说这是一个好策略。寻找好策略不关我的事；这事归遗传算法管。

我写了一个遗传算法程序来进化罗比的策略。算法中，群体中每个个体都是一个策略——与各种可能情形相对应的行动列表。也就是说，对于表9—1中的策略，GA用来演化的个体就是最右侧243个行动依次列出的列表：

向北移动　　向东移动　　随机移动　　收集罐子……向西移动……不动

字符串中第1个行动（这里是向北移动）对应第1种情形（“空空空空空”），第2个行动（这里是向东移动）对应第2种情形（“空空空空罐”），依次往后。这样就不用明确列出与动作对应的情形；GA记得各情形的排列顺序。例如，假设罗比观察到情形如下：

北	南	东	西	中
空	空	空	空	罐

GA根据内建知识知道是情形2。通过查询策略表可以得知位置2上的行动是往东移动。罗比往东移动一格，然后又观察周围情形；GA再次查询表上的相应行动，反复进行。

我的GA是用C语言写的。这里我不写出具体程序，只解释其工作原理。

1.生成初始群体。初始群体有200个随机个体（策略）。

图9.3是一个随机群体示意图。每个个体策略有243个“基因”。每个基因是一个介于0和6之间的数字，代表一次动作（0=向北移动，1=向南移动，2=向东移动，3=向西移动，4=不动，5=捡拾罐子， 6=随机移动）。在初始群体中，基因都随机设定。程序中用一个伪随机数发生器来进行各种随机选择。

重复后面的步骤1000次。

个体 1:
2330032342163034353054600610256251511416226043565433406651151415650220640642051006643216161521652022364433363346013326503000406220502431650061113051466642324012456333455241261434413610201506306425516540432644631561645105436653463105516460051 64

个体 2:
1641134312102536034036124143120110423546252530420204451643366561035322153105131440622120614631432154610256523644422025340345305020056206340263310024534164301516312100122144006640126652463516501541231131324533044332126345550053142130644233110 00

个体 3:
2042334440241122613213645263246421220612212225266062614443612532512664061335340153411110206164226653145522540234051155031302220200654451250622066314261355320100004000316401301541601620061344406261605056414215531332360215033551312536326426305 51

.

.

.

个体 200:
3463252513600101222561210604330113520515532013065600532223504332425064124255265534635345523053326612010632124554423440613654302462401606630164646411030265400063341261503522621060636242605506166163442551243544641100234633304401025332121424022 51

图9.3　随机初始群体。每个个体由243个数字组成，取值介于0到6之间，每个数字编码一个动作。数字的位置决定它对应哪种情形

2.计算群体中每个个体的适应度。在我的程序中，是通过让罗比执行100次不同的清扫任务来确定策略的适应度。每次将罗比置于位置（0，0），随机撒一些易拉罐（每个格子至多1个易拉罐，格子有易拉罐的概率是50%）。然后让罗比根据策略在每次任务中执行200个动作。罗比的得分就是策略执行各任务的分数。策略的适应度是执行100次任务的平均得分，每次的罐子分布都不一样。

3.进化。让当前群体进化，产生出下一代群体。即重复以下步骤，直到新群体有200个个体。

（a）根据适应度随机选择出一对个体A和B作为父母。策略的适应度越高，被选中的概率则越大。

（b）父母交配产生两个子代个体。随机选择一个位置将两个数字串截断；将A的前段与B的后段合在一起形成一个子代个体，将A的后段与B的前段个体合在一起形成另一个子代个体。

（c）让子代个体以很小的概率产生变异。以小概率选出1个或几个数，用0到6的随机数替换。

（d）将产生的两个子代个体放入新群体中。

4.新群体产生200个个体后，回到第2步，对新一代群体进行处理。

神奇的是，从200个随机的策略出发，遗传算法就能产生出让罗比顺利执行任务的策略。

种群规模（200）、迭代次数（1000）、罗比在一次任务中的动作数量（200）以及计算适应度的任务数量（200）都是我设定的，有些随意。用其他参数也能产生出好的策略。

你现在肯定很想知道遗传算法得出的结果。不过，我得向你坦白，在运行程序之前我还是克服了懒惰，自己设计了一个“聪明的”策略，这样我就能知道GA和我比起来谁干得更好。我为罗比设计的策略是：“如果当前位置有罐子，就捡起来。否则，如果旁边格子有罐子，就移过去。（如果有几个罐子，预先设定罗比向哪个移动。）否则，随机选择一个方向移动。”

这个策略其实不是很聪明，罗比有可能会围着空格子绕圈子，总也找不到易拉罐。

我用10000个清扫任务测试了这个策略，结果（每个任务的）平均分约为346。每个任务最初有大约50%的格子有罐子，也就是50个易拉罐，因此最高可能的分数约为500，这样看我的策略还不是很接近最

优。

GA能有这么好的成绩吗？会不会更好？运行一下就知道了。取最后一代中适应度最高的个体，也用10000个不同的任务进行测试。结果平均分约为483——几乎是最优了！

GA演化的策略是如何解决这个问题的

问题是这个策略是如何做到的？它为什么能比我的策略做得更好？GA又是如何将它演化出来的？

将我的策略记为M， GA生成的策略记为G。下面是这两个策略的基因组。

M：656353656252353252656353656151353151252353252151353151656353656252353252656353656050353050252353252050353050151353151252353252151353151050353050252353252050353050656353656252353252656353656151353151252353252151353151656353656252353252656353454

G：254355153256235251056355461151336154151034156110550150052030256256132252350325112052333054055231255051336154150665264150266506012264453605631520256431054354632404350334153250253251352352045150130156213436252353223135051260513356201524514343432

仅仅从策略的基因组看不出其中的运作。我们可以知道其中一些基因的意义，例如当前位置有罐子的情形，对第2种情形（“空空空空罐”），两个策略的动作都是5（清扫罐子）。M对这种情形总是动作5，但G并不总是这样。例如下面这种情形：

北	南	东	西	中
空	罐	空	罐	罐

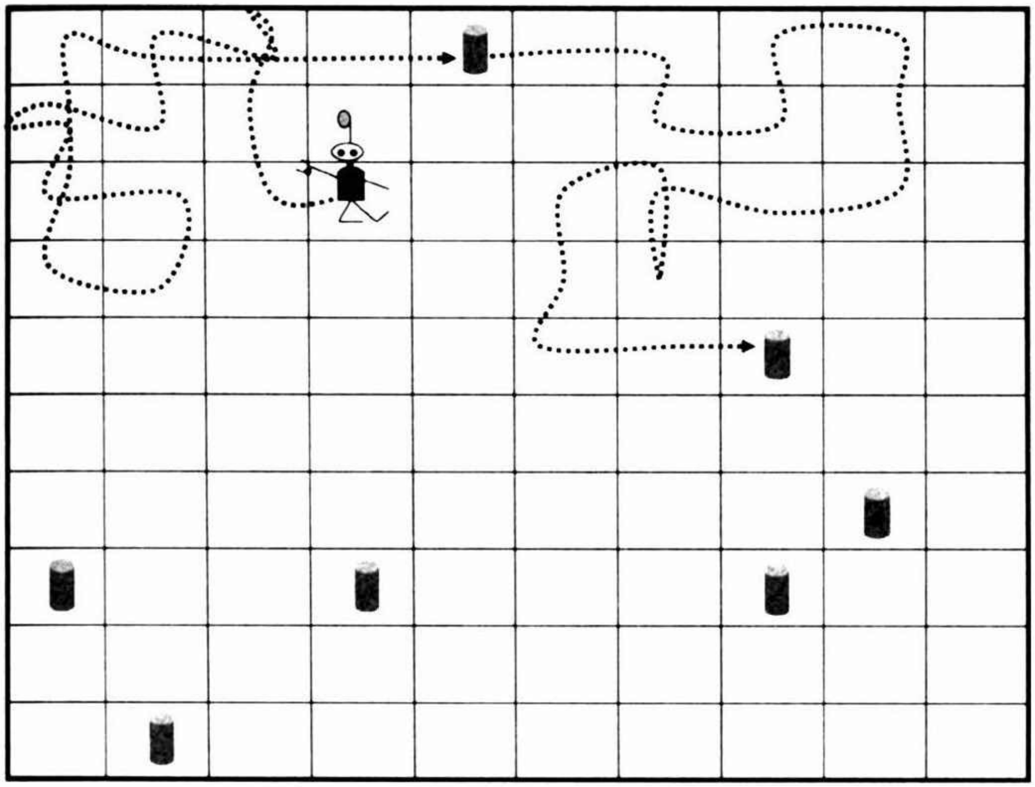
动作就是3（往西移动），罗比在这种情形下不会捡罐子。这不太好，不过G总体上还是比M好。

关键之处不在于单个的基因，而在于各个基因之间的相互作用，就像真正的基因一样。而且同真正的基因一样，很难确定各种相互作用是如何影响整体上的行为或适应度。

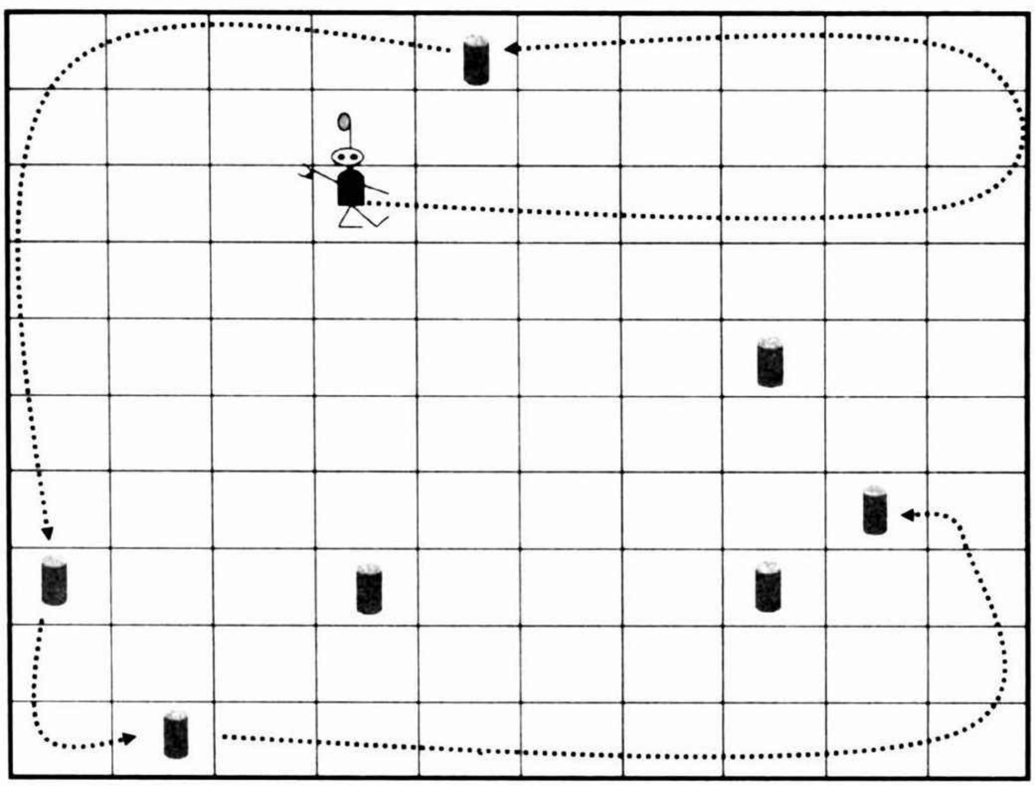
相较于观察一个策略的基因，观察其具体的行为——也就是它们的表型——会更有意义。我编了一个程序来演示罗比在采用某个给定策略时的行动，然后对罗比在采用策略M和策略G时的行为进行观察。我发现这两个策略在许多情形中的行为类似，但策略G有两个小技巧，让它比策略M表现得更好。

首先来看看当前位置和四周都没有罐子的情形。如果罗比采用策略M，它就会随机选择一个方向移动。但如果它采用的是策略G，它就会往东移动，直到遇到墙为止。然后它会往北移动，就这样逆时针围着格子边缘移动，直到发现罐子。图9.4可以看到罗比的轨迹（虚线）。

这种围着绕圈的策略不仅让罗比不会撞墙（如果用策略M，随机移动时有可能会撞墙），而且搜索罐子的效率也比随机移动要高。



策略M



策略G

图9.4　罗比在一个有罐子的场地中。图中虚线是它在仿真时的移动轨迹，上图采用策略M，下图采用策略G

另外，遗传算法通过策略G还发现了一个巧妙的技巧，它在一些特定的情形中不会去捡当前格子中的罐子。

例如图9.5（a）给出了一种情形。在这种情形下，如果罗比采用策略M，它会捡起当前格子里的罐子，向西移动，然后捡起新的格子里的罐子[图9.5（b）—图9.5（d）]。由于罗比只能看见相邻格子的情形，因此当前它看不到余下的一堆罐子。它只能随机移动，直到碰巧遇到余下的罐子。

再看看策略G在同样情形中的表现（图9.6）。罗比没有去捡当前位置上的罐子，而且是直接向西移动[图9.6（b）]。然后它捡起了一堆罐子中最西边的罐子[图9.6（c）]。前面没有捡的罐子现在成了路标，罗比根据这个可以“记住”返回去有罐子。接下来它就会把这一堆罐子都捡起来[图9.6（d）—图9.6（k）]。

我知道我的策略不完美，但也没想到会有这种办法。进化可能聪明得多，GA经常会让我们感到意外。

遗传学家经常用“敲除突变（knockout mutations）”的办法来验证关于基因功能的理论，其实就是用遗传工程的方法阻止所研究的基因转录，再看对机体会有何影响。这里我也可以用这种方法。我将策略G中与这个技巧对应的基因敲掉：将所有与“当前格子中有罐子的”情形相对应的基因都换成“清扫罐子”。这会使得策略G的平均分从最初的483降到443，因此证明了我在前面的猜测，策略G之所以成功，部分就是因为这个技巧。

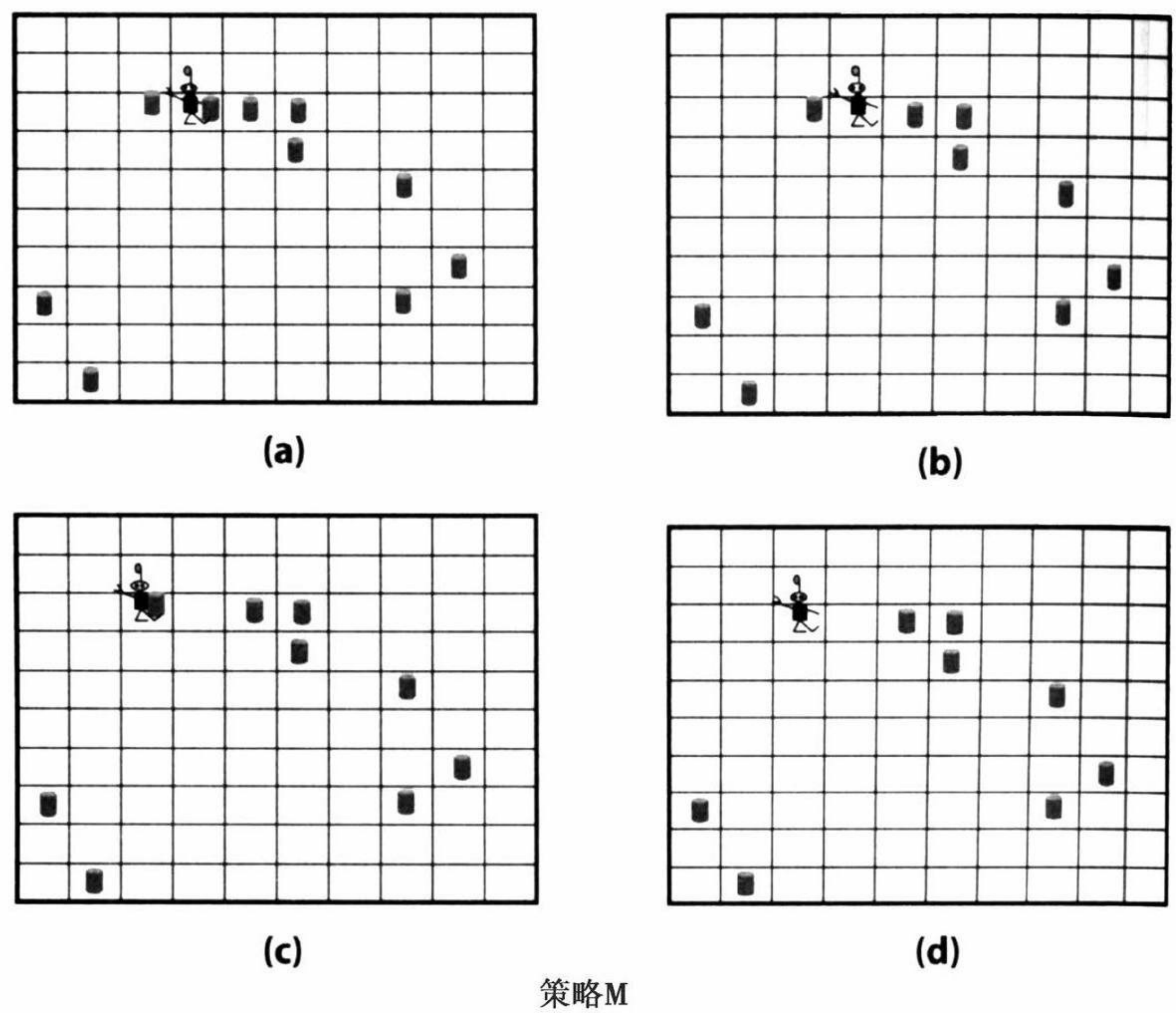
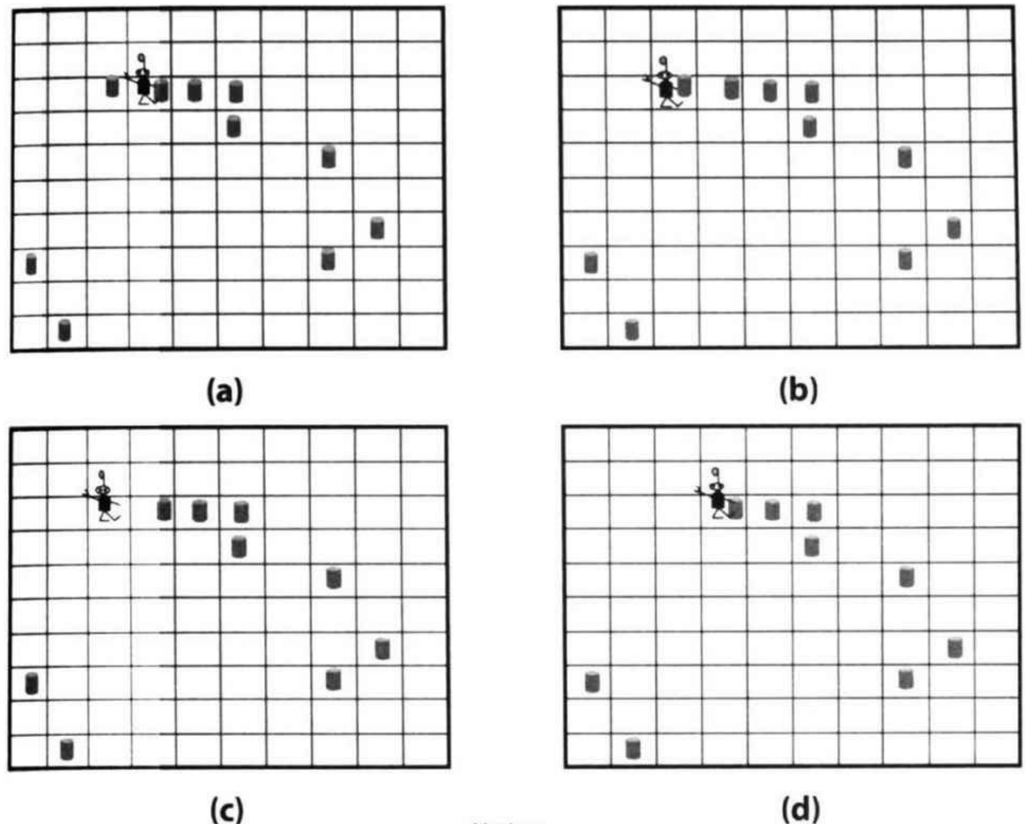
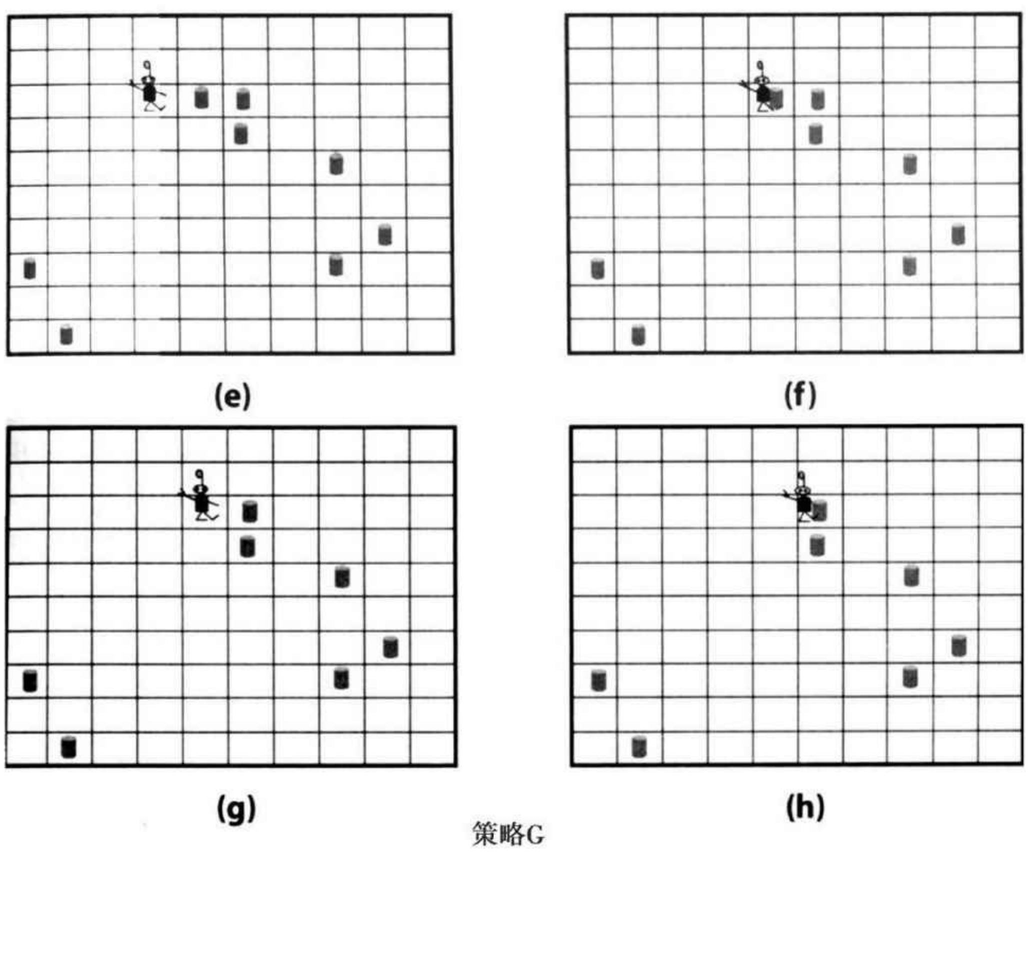


图9.5　罗比在一堆罐子中，使用策略M移动的四步



(a) (b) (c) (d)

策略G



(e) (f) (g) (h)

策略G

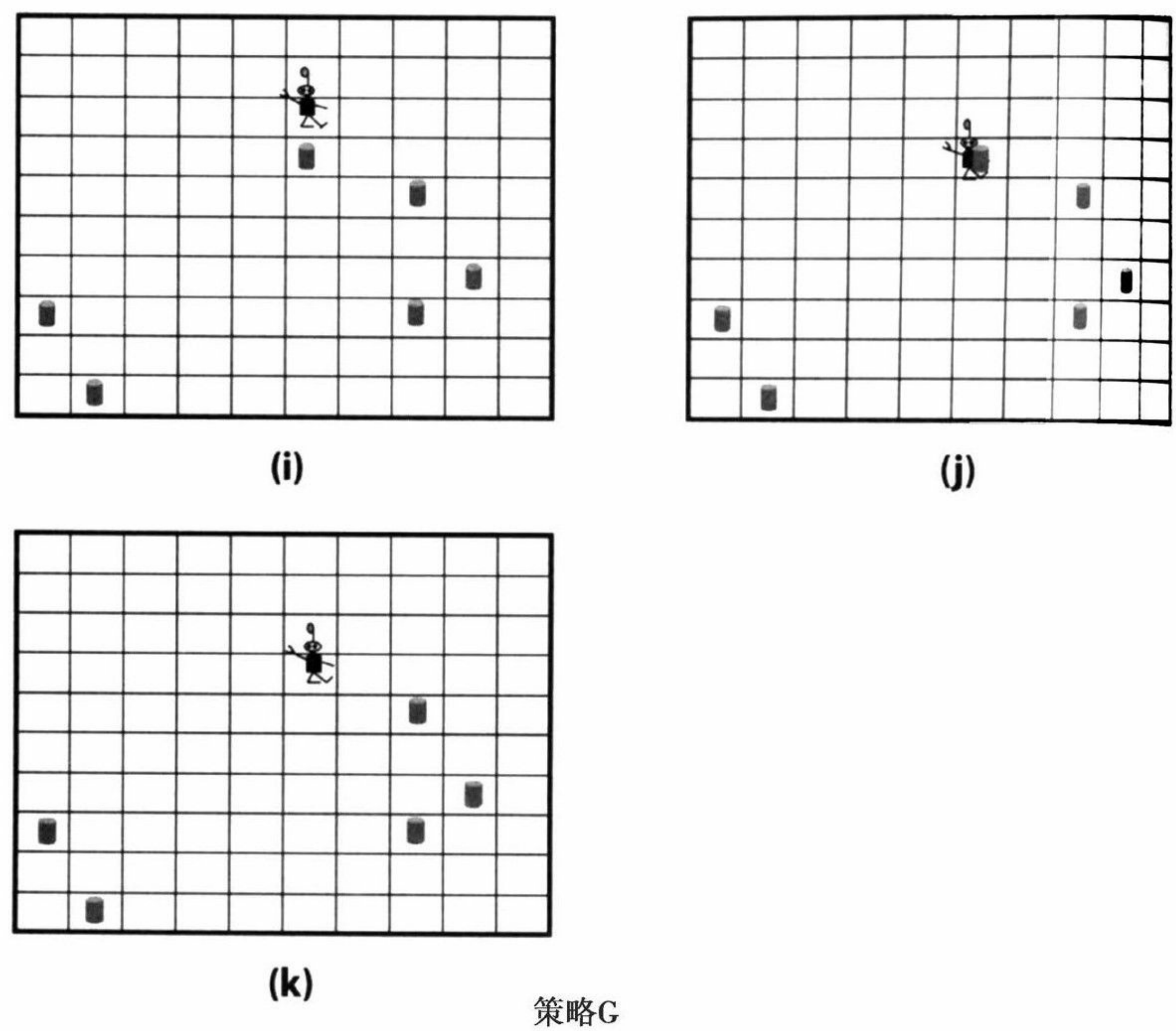


图9.6　罗比在同样的一堆罐子中，用策略G走的11步

GA是如何演化出好的技巧的

下一个问题是，GA是如何从随机的群体演化出像策略G这样好的策略的呢？

要回答这个问题，我们可以看一看策略是如何一代一代改进的。图9.7中画出了每一代中最佳策略的适应度。你可以看到最好的适应度最开始是小于0的，前300代提高得很快，此后的提高要慢一些。

第1代有200个随机生成的策略，可以想象它们都很糟糕。最好的策略适应度才-81，最糟糕的到了-825。（可能这么低吗？）

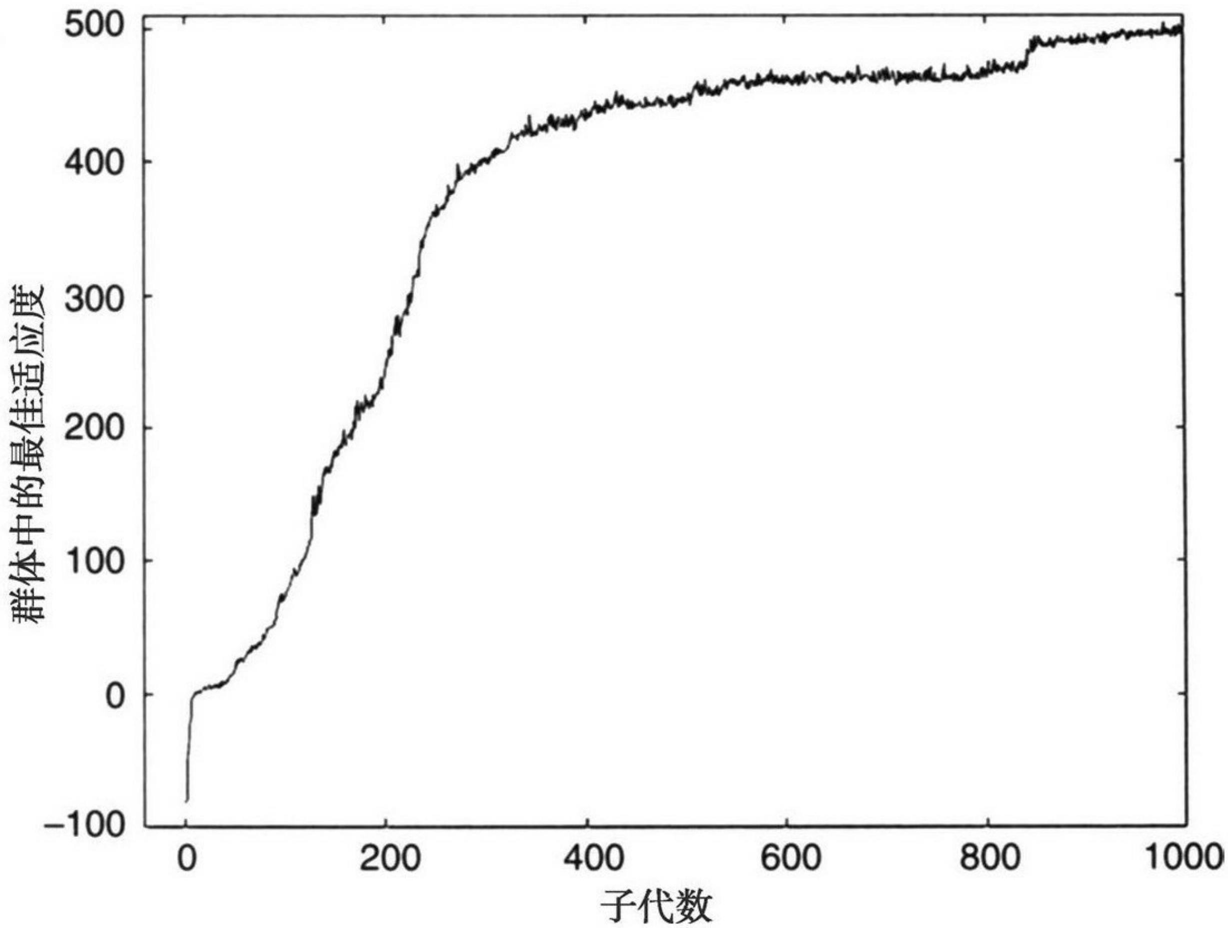


图9.7　GA演化出策略G的过程中，各代群体中的最佳适应度

我用几个任务测试了一下罗比采用这一代中最糟糕的策略时的行为。在一些环境设定中，罗比移动了几步就卡住了，之后在整个任务过程中都停止不动。在一些情况下，则不停地撞墙，直到任务结束。有时候则一直不断地去捡罐子，虽然当前位置上没有罐子。显然这些策略在进化过程中很快就会被淘汰掉。

我也测试了一下这一代中最好的策略，还是很糟糕，比最差的好不了多少。不过比起来它还是有两个优点：不那么容易一直撞墙了，而且偶尔碰到罐子的时候还能把罐子捡起来！作为这一代中最好的策略，它有很大的机会被选中用来繁殖！一旦被选中，它的子代就会继承这些优点（同时也会继承许多缺点）。

到第10代，群体中最佳策略的适应度已经变成正数了。这个策略经常会停滞不动，有时候还会在两个格子之间不停地来回移动。但基本不怎么撞墙，同第1代的前辈一样，偶尔也会捡罐子。

GA就这样不断改进最佳适应度。到200代时，最好的策略已经具有向罐子移动并捡起罐子这个最重要的能力——至少大部分时候是这样。不过，如果周围没有罐子，它也会浪费很多时间用来随机游走，这一点同策略M相似。到250代时，做得已经和策略M一样好了；等到了400代，适应度超过了400分，这时的策略如果少做一些随机移动，就能和策略G一样好。到800代时，GA发现了将罐子留作相邻罐子的路标的技巧，到900代时，沿着围墙转的技巧就基本完善了，到1000代时会进一步做些修正。

虽然罗比机器人的例子相当简单，但它与实际应用的GA区别已不是很大。同罗比的例子一样，在实际应用中，GA经常能演化出有用的答案，但是很难看出为什么会有用。这是因为GA找到的好答案与人类想出的相当不同。美国国家航空航天局（NASA）的遗传算法专家罗恩（Jason Lohn）曾这样说：“进化算法是探索设计死角的伟大工具。[143]你向具有25年工业经验的专家展示（你的设计），他们会说‘哦，这个真的能有效？’……我们经常发现进化出来的设计完全无法理解。”

罗恩的设计也许是无法理解，但的确能有效。2004年，罗恩和他的同事因为用GA设计出了新的NASA航天器天线被授予“人类竞争”奖（Human Competitive Award）。这表明GA的设计改进了人类工程师的设计。

3　大写的计算

计算机科学的真义[144]是在自然界中无处不在的大写的计算。

——朗顿（Chris Langton），

引自勒温（Roger Lewin）的

《复杂性：混沌边缘的生命》

（Complexity: Life at the Edge of Chaos ）

第10章　元胞自动机、生命和宇宙

自然界中的计算

《科学》杂志不久前出版了一篇文章，[145]题为《社会性昆虫行为的计算》（*Getting the Behavior of Social Insects to Compute* ），文中介绍了一些昆虫学家的工作，他们将蚂蚁群体的行为等同于“计算机算法”，每只蚂蚁都执行简单的程序，使得整个种群作为一个整体执行复杂的计算，比如在决定何时将巢穴搬往何地的问题上形成一致。

如果由一只蚂蚁来领导和决策，很容易就能在计算机上编出程序来进行计算。其他蚂蚁只需按照领导者的决策来做就行了。然而，在前面我们已经看到，在蚂蚁群体中没有领导者；蚁群“计算机”由数百万只自主的蚂蚁组成，每只蚂蚁都只是根据与一小部分蚂蚁的交互来进行决策和行动。这种计算与具有CPU和内存的普通电脑进行的计算差别很大。

与此类似，1994年三位杰出的大脑科学家也写了一篇文章：“大脑是计算机吗？”[146]他们认为：“如果我们同意接受更宽泛的计算概念，答案就一定是‘是’。”同蚁群一样，大脑的计算方式——数以亿计的神经元并行工作，而无须中央控制——也与现代的数字计算机的运作方式完

全不同。

在前两章我们探讨了计算机中的生命和进化。在这一部分，我们来看看相对的思想，以及计算在自然界中的广泛存在。自然系统的“计算”指的是什么呢？大致上说，计算是复杂系统为了成功适应环境而对信息的处理。但是这样的说法还能更精确些吗？信息在哪里？复杂系统又是如何处理信息的？

为了让这类问题更易于研究，科学家们通常会将问题理想化——也就是尽可能简化，但仍然保留问题的主要特征。鉴于此，许多人都用元胞自动机这种理想化的复杂系统模型来研究自然界中的计算。

元胞自动机

在第4章曾讲过，图灵机将“明确程序”——也就是计算——的概念进行了形式化。计算就是图灵机根据机器的规则集将带子上的初始输入转换成停机时带子上的输出。这个抽象的机器就是后来所有数字计算机的设计原型。由于冯·诺依曼对计算机设计作出的贡献，现在的计算机架构被称为“冯·诺依曼体系结构”。

冯·诺依曼体系结构包括存储数据和程序指令的随机存取存储器（RAM）、从存储器存取指令和数据并执行指令处理数据的中央处理单元（CPU）。你可能知道，虽然程序员们编程时使用的是高级语言，存储在计算机中的指令和数据却是0/1组成的串。执行指令就是将这种0/1码译成基本的逻辑操作让CPU执行。只需要几种基本的逻辑操作就能实现所有计算，现代CPU每秒能执行数亿次这样的逻辑操作。

元胞自动机是理想化的复杂系统，结构完全不同于计算机。想象一块板子上排列着许多灯泡（图10.1），每个灯泡与四周以及斜对角的灯泡连在一起。在图中只画了其中一个灯泡的连接线，不过姑且想象所有灯泡都有连接线。

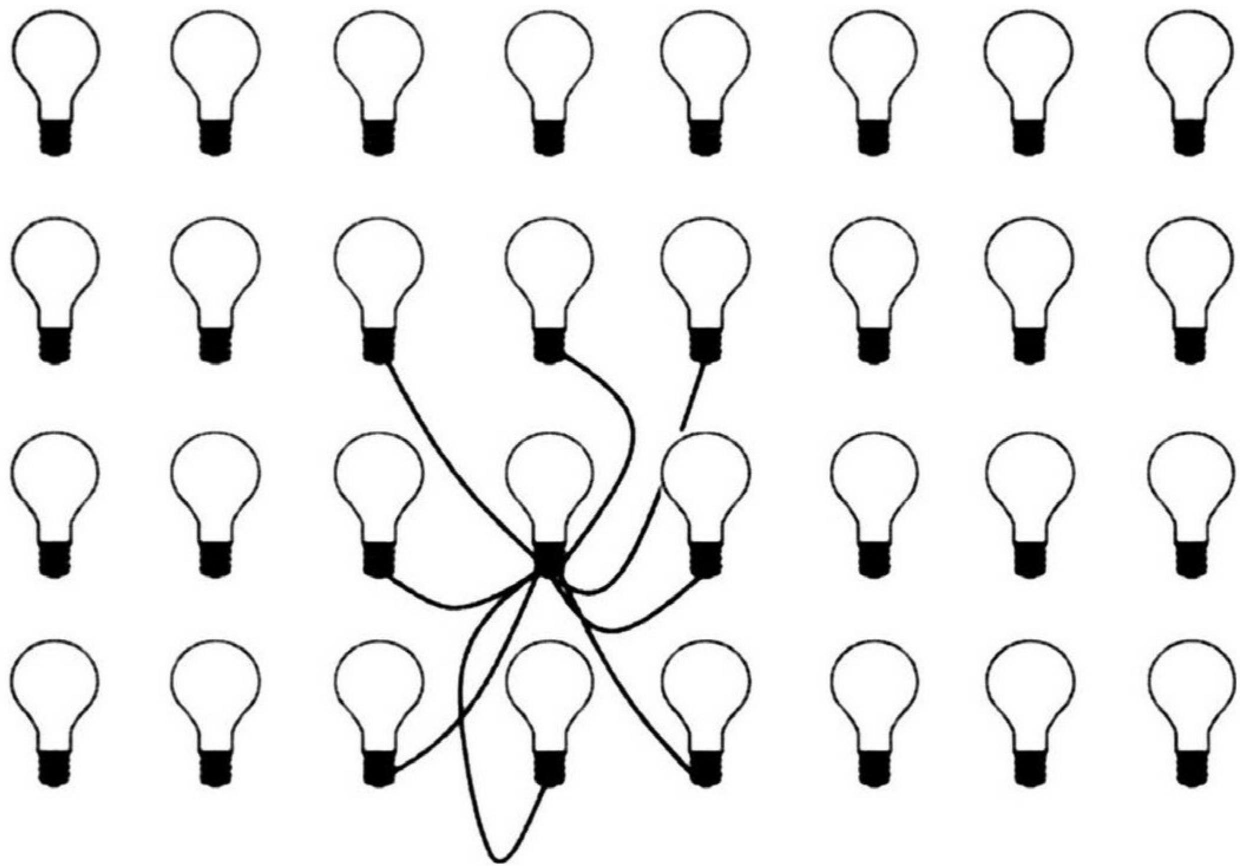


图10.1　排列的灯泡，每个都与四周以及对角的灯泡相连，图中画了一个灯泡的连线作为示范。灯泡的状态可以是亮和灭。假设边沿是回绕连在一起，也就是认为最左边的与最右边的灯泡相邻，最下面的与最上面的灯泡相邻，等等

图10.2中（左边的盒子），有些灯泡已经点亮（为了简洁，我没有画灯泡的连线）。先设定好灯泡的开关状态，然后各个灯泡开始不断定时“更新状态”——选择开或关，所有灯泡都同步变化。你可以将这个灯泡阵看作萤火虫发光的模型，每只萤火虫都根据周围萤火虫的闪灭来调整自己是亮还是灭；也可以看作神经元的激发模型，各个神经元受周围神经元的状态激发或抑制；或者就当作抽象艺术也行，如果你愿意的话。

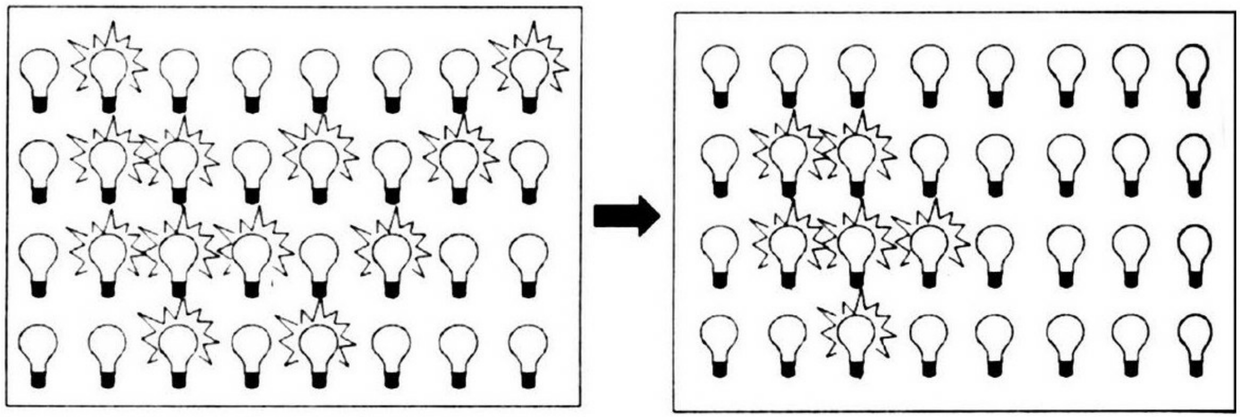


图10.2　左：灯泡阵列的初始状态，没有画灯泡之间的连线。右：变化一次之后的状态，规则是“采用邻域占多数的状态”

灯泡每一步如何“决定”是开还是关呢？它们都遵循一些规则，根据邻域内灯泡的状态——也就是相邻的8个灯泡和它自己的状态——来决定下一步的状态（是开还是关）。

例如，规则可以是这样：“如果邻域内的灯泡（包括自己）点亮的超过一半，就点亮（如果本来就是亮的，则不变），否则就熄灭（如果本来就是灭的，则不变）。”也就是说，邻域中9个灯泡，如果有5个或5个以上是亮的，中间的灯泡下一步就是亮的。我们来看看灯泡阵列下一步会怎么变。

图10.1的文字中说了，为了让每个灯泡都有8个邻居，阵列的四边是回绕相连的。可以想象成上边和下边合到一起，左边和右边合到一起，形成一个面包圈形状。这样每个灯泡就都有8个邻居。

现在再来看上面给出的规则。图10.2是初始设置和变化一次后的状态。

也可以使用更复杂的规则，例如，“如果邻域中点亮的灯泡不少于2个，不多于7个，就点亮，否则就熄灭”，这样阵列的变化就会不一样。或者这样，“如果刚好1个灯泡是灭的，或者4个是亮的，就点亮，否则就熄灭”。可能的规则很多。

到底有多少种可能的规则呢？说“很多”还太保守了。答案是“2的512次幂”（2^{512}）[147]，这个数字很大，比宇宙中的原子数量还大许多倍。（注释中有答案的推导过程。）

这个灯泡阵列其实就是一个元胞自动机。元胞自动机是由元胞组成的网格，每个元胞都根据邻域的状态来选择开或关。（广义上，元胞的状态可以随便定多少种，但是这里我们只讨论开/关状态。）所有的元胞遵循同样的规则，也称为元胞的更新规则，规则根据各元胞邻域的当前状态决定元胞的下一步状态。

为什么说这么简单的系统会是复杂系统的理想化模型呢？同自然界的复杂系统一样，元胞自动机也是由大量简单个体（元胞）组成，不存在中央控制，每个个体都只与少量其他个体交互。而且元胞自动机也能表现出非常复杂的行为，它们的行为很难甚至不可能通过其更新规则来预测。

同其他许多精彩的思想一样，元胞自动机也是由冯·诺依曼发明的，他在20世纪40年代受他的一位同事——数学家乌拉姆——的启发提出了这个思想。（为了与冯·诺依曼体系结构相区别，元胞自动机经常被称为非冯·诺依曼体系结构，这是计算机科学的一大笑话。）第8章说过，冯·诺依曼想要将自我复制机器的逻辑形式化，而他用来研究这个问题的工具就是元胞自动机。简单地说，他设计了一种元胞自动机规则，能完美复制任意元胞自动机的初始形态，他的规则中元胞不止两种状态，而是29种。

冯·诺依曼还证明他的元胞自动机等价于通用图灵机（参见第4章）。元胞的更新规则扮演了图灵机读写头的规则的角色，而元胞阵列的状态则相当于图灵机的带子——也就是说，它可以编码通用图灵机运行的程序和数据。元胞一步一步地更新相当于通用图灵机一步一步地迭代。能力等价于通用图灵机的系统（也就是说，通用图灵机能做的，它也能做）被称为通用计算机，或者说能进行通用计算。

生命游戏

冯·诺依曼的元胞自动机规则相当复杂。1970年，数学家康威（John Conway）发现了一种简单得多的两状态通用图灵机，也能进行通用计算。他称之为“生命游戏”。[148] 我不知道为什么叫作“游戏”，只知道“生命”来自于康威对其规则的解释。将状态为开的元胞看作活的，状态为关的元胞看作死的。康威用四种生命过程来定义规则：出生，死元胞的相邻元胞中如果刚好有3个是活的，下一步就变成活的；存活，

活元胞的相邻元胞有2—3个是活的，下一步就能继续存活；过于稀疏，活元胞的相邻活元胞如果少于2个就会死去；过度拥挤，活元胞的相邻活元胞如果多于3个就会死去。

康威当时是想寻找一个能产生类似生命的元胞自动机。出人意料的是，生命游戏的行为丰富而有趣，以至于现在出现了一个爱好者团体，他们的主要兴趣就是发现能产生有趣行为的初始设置。

图10.3就是其中一个有趣的行为，被称为滑翔机。这里我们不再用灯泡，就用黑格子表示开（活），用白格子表示关（死）。图10.3中可以看到有一个滑翔机向东南方向移动。当然，元胞并没有动，它们都是固定的。移动的是由活状态元胞形成的一个不消散的形状。因为元胞自动机的边界是回绕相连的，所以滑翔机能一直移动。

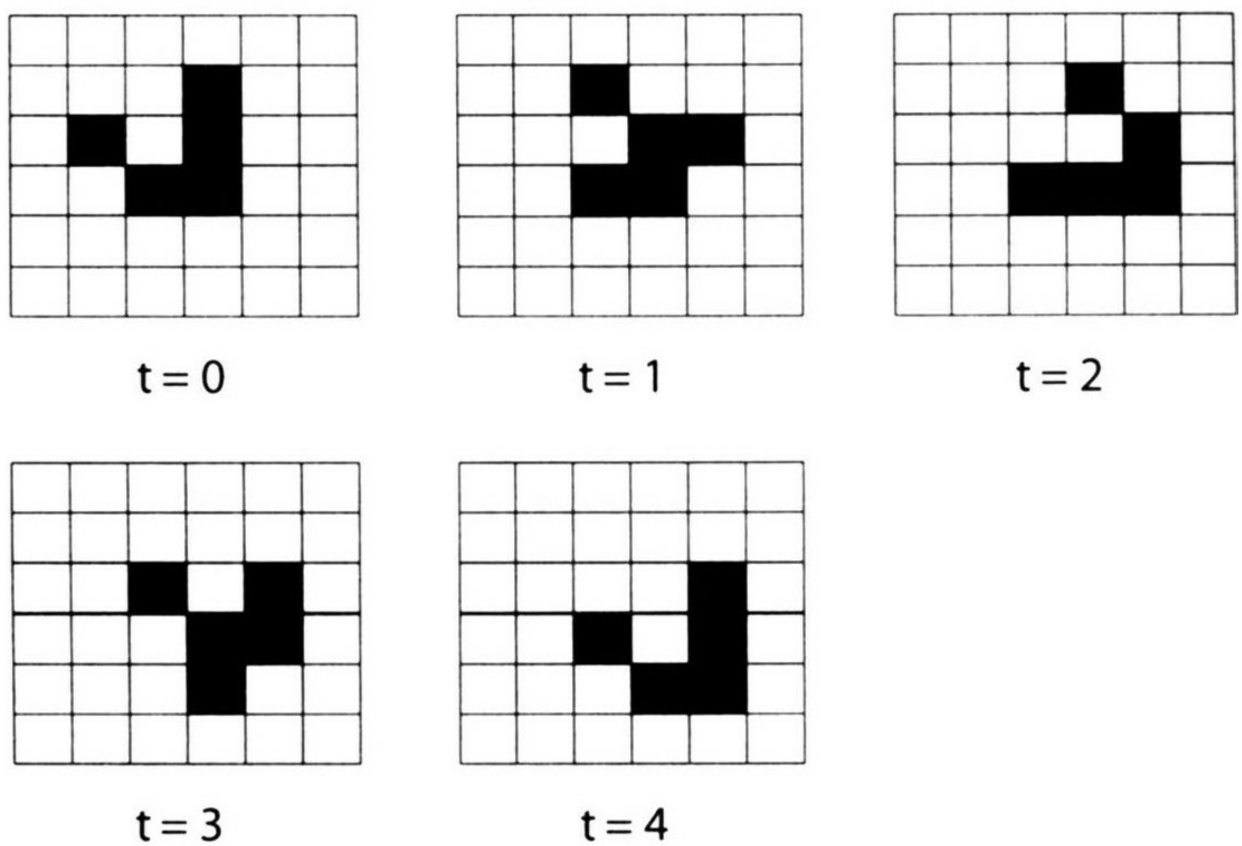


图10.3　生命游戏中滑翔机的一个循环变化。滑翔机形状每4步向东南方向移动一格

爱好者们还发现了其他复杂的形状，例如太空船，它类似于滑翔机，而且更有趣；滑翔机发射器，它会不断射出新的滑翔机。康威证

明，通过用变化的开/关状态模拟读写头在带子上的读写，就能让生命游戏模拟图灵机。

康威还给出了生命游戏模拟通用计算机的证明框架[149]（后来由其他人细化）。[150]将程序和输入数据编码为开关状态初始设置，生命游戏运行后产生的图样表示程序的输出。

康威在证明中组合使用滑翔机发射器、滑翔机等结构来实现与、或、非等逻辑运算。人们早已发现，能用各种可能来组合这些逻辑操作的机器就能进行通用计算。康威的证明表明，原则上，逻辑运算的所有可能组合都能在生命游戏中实现。

像生命游戏这么简单的元胞自动机在原则上能运行标准计算机运行的程序，这真是让人吃惊。不过，实际上，稍微复杂一点的计算就需要大量逻辑运算，并以各种方式相互作用，因此要设计出能实现复杂计算的初始设置基本不太可能。即使设计得出来，计算也会慢得让人无法忍受，更不要说用这种并行的非冯·诺依曼结构的元胞自动机来模拟传统冯·诺依曼结构计算机需要耗费的大量资源。

因此没有人用生命游戏（或其他“通用”元胞自动机）来进行真实计算或是模拟自然系统。我们只是想利用元胞自动机的并行特征以及它产生复杂图形的能力。下面先来看看元胞自动机能够产生的图形种类。

四类元胞机

20世纪80年代初，普林斯顿高等研究院的物理学家沃尔夫勒姆对元胞机着了迷。沃尔夫勒姆（图10.4）是传奇般的天才人物。他1959年生于伦敦，15岁就发表了他的第一篇物理学论文。两年后，在牛津大学一年级的暑假，大部分人这时候都会去打工挣钱，或是背着背包搭顺风车周游欧洲，沃尔夫勒姆却写了一篇关于“量子色动力学”的论文，这篇论文被诺贝尔奖获得者物理学家盖尔曼注意到，他邀请沃尔夫勒姆加入他在加州理工的研究小组。两年后，沃尔夫勒姆获得了理论物理博士学位，而这时他才20岁（大部分人大学毕业后至少需要5年才能获得博士学位）。他留在加州理工任教，此后不久又获得了第一届麦克阿瑟天才奖。两年后，他受邀加入普林斯顿高等研究院。真是让人叹为观止。他有了名气，又有资金支持，可以想干什么就干什么，他决定研究元胞自

动机动力学。

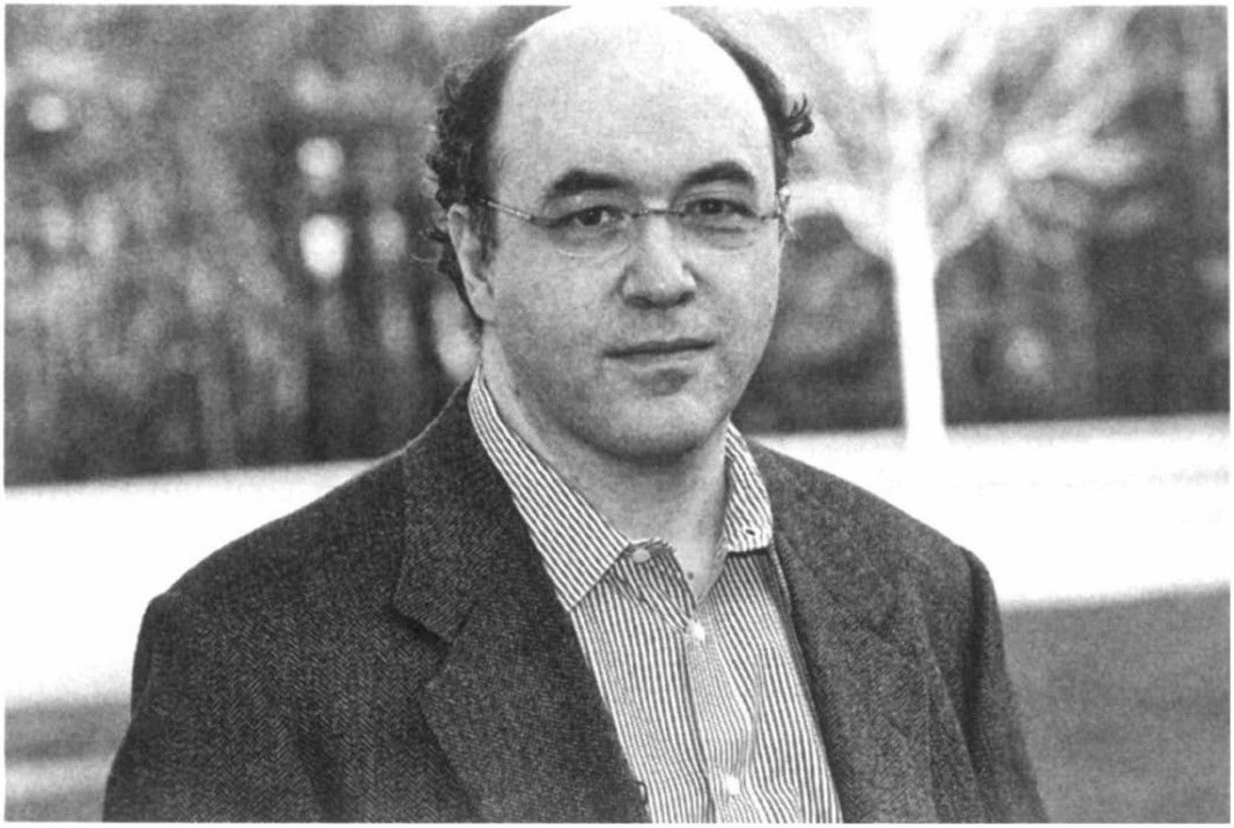


图10.4　沃尔夫勒姆（照片由沃尔夫勒姆研究公司提供，Wolfram Research, Inc.）

根据理论物理学的惯例，沃尔夫勒姆从元胞自动机最简单的形式来研究其行为，他用的是一维两状态的元胞自动机，每个元胞仅与两个相邻元胞相连[图10.5（a）]。沃尔夫勒姆称之为“初等元胞自动机（elementary cellular automata）”。他认为，如果这种看上去极为简单的系统也理解不了，就更不可能理解更复杂的（例如，两维或多状态的）元胞自动机。

图10.5描绘了一个初等元胞自动机的规则。图10.5（a）是元胞格子——一排元胞，每个都与两侧相邻的元胞相连。这里仍然是用方格表示元胞——黑表示开，白表示关。两头的格子回绕连在一起，形成一个环。图10.5（b）是元胞遵循的规则：3个相邻元胞总共有8种可能状态组合，对于每种状态组合都给出了中间元胞的更新状态。例如，当3个元胞都是关状态时，中间元胞下一步就是关状态。同样，当3个元胞的状态为关—关—开时，中间元胞下一步就会变成开状态。这里规则指的是整个状态组合和更新状态的表，而不仅仅是表中的一行。图10.5（c）

展示了这个元胞自动机的时空图。最顶上一行是一维元胞机的初始状态设置。下面跟着的依次是每一步更新后的状态。这种图被称为时空图，因为它表现了元胞自动机的立体构型随时间的变化。

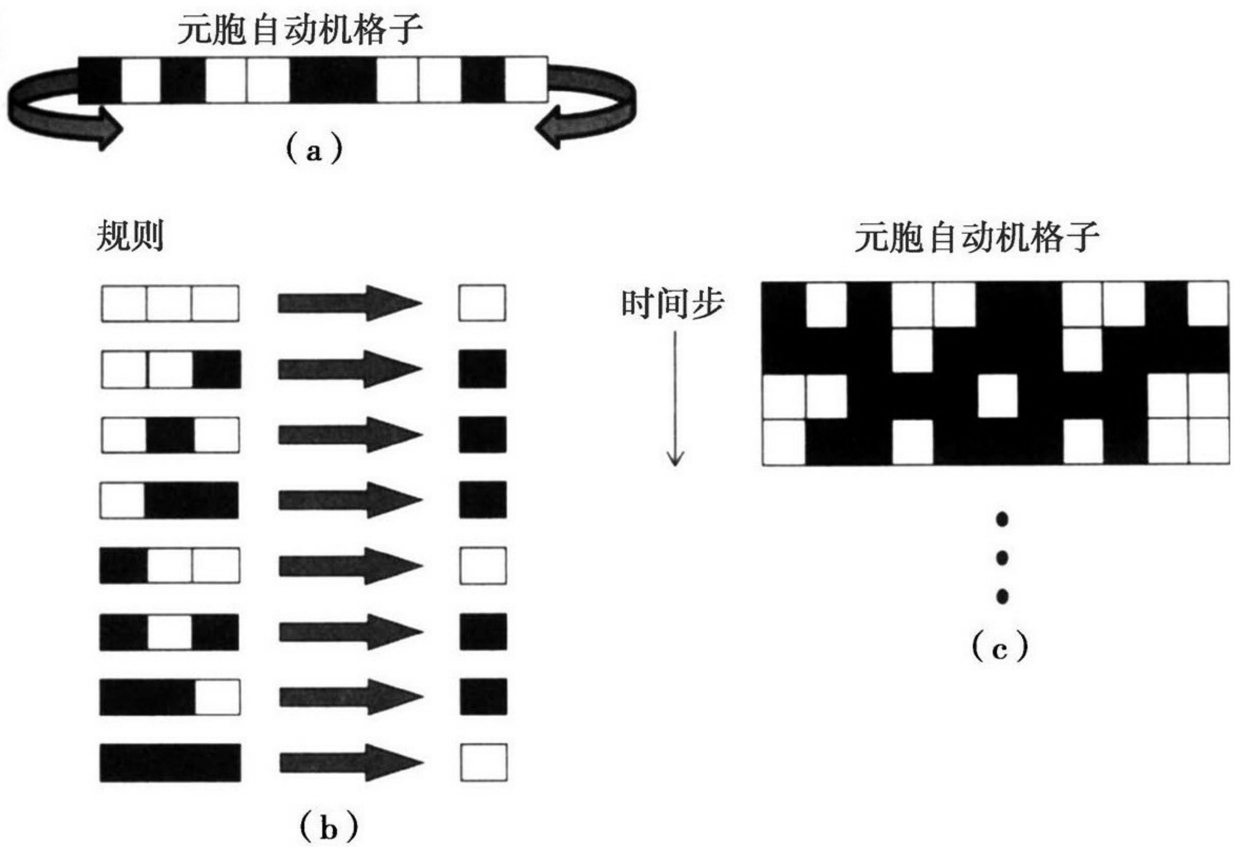


图10.5　（a）一维元胞自动机，两端回绕形成一个环；（b）初等元胞自动机的一种规则（规则110），表示3个元胞的状态组合以及对应的中间元胞下一步的状态；（c）时空图，显示了元胞自动机的4步变化

3个元胞有8种可能的状态组合[图10.5（b）]，而每种可能的状态组合又有两种可能的元胞更新状态，因此初等元胞自动机总共有256（2^8）种可能的规则。20世纪80年代时，计算机的性能已经足以让沃尔夫勒姆对这些规则逐个进行研究，设定各种初始状态，然后观察它们的变化。

沃尔夫勒姆给初等元胞自动机的规则都编了号，编号方法见图10.6。他将开状态记为“1”，关状态记为“0”，根据更新状态将规则记为0/1串，最前面一位对应3个元胞都为开时的更新状态，最后一位则对应3个元胞都为关时的更新状态。这样图10.5中的规则就记为0 1 1 0 1 1 1 0。然后沃尔夫勒姆将这个0/1串当作二进制数。0 1 1 0 1 1 1 0作为二进

制数等于十进制数110。因此这个规则就叫作“规则110”。如果更新状态是0 0 0 1 1 1 1 0，则为“规则30”。（注释中介绍了如何将二进制数转换为十进制数。[151]）

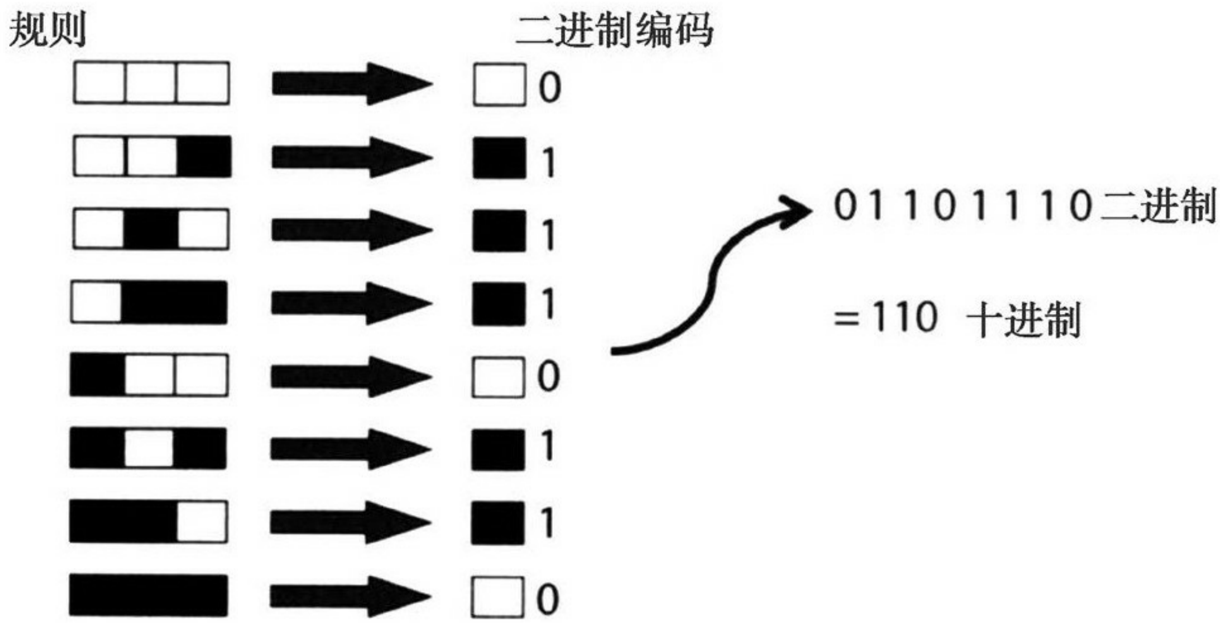


图10.6　沃尔夫勒姆使用的初等元胞自动机命名系统

沃尔夫勒姆和他的同事开发了一种专门的计算机语言——Mathematica——用来简化元胞自动机的模拟。沃尔夫勒姆用Mathematica编程运行元胞自动机，并绘制它们的时空图。例如，图10.7和图10.8与图10.5类似，只是规模大些。图10.7中有200个元胞，初始状态随机设置，应用的更新规则是110，逐行往下更新了200时间步。图10.8应用的则是规则30，也是从随机初始状态开始。

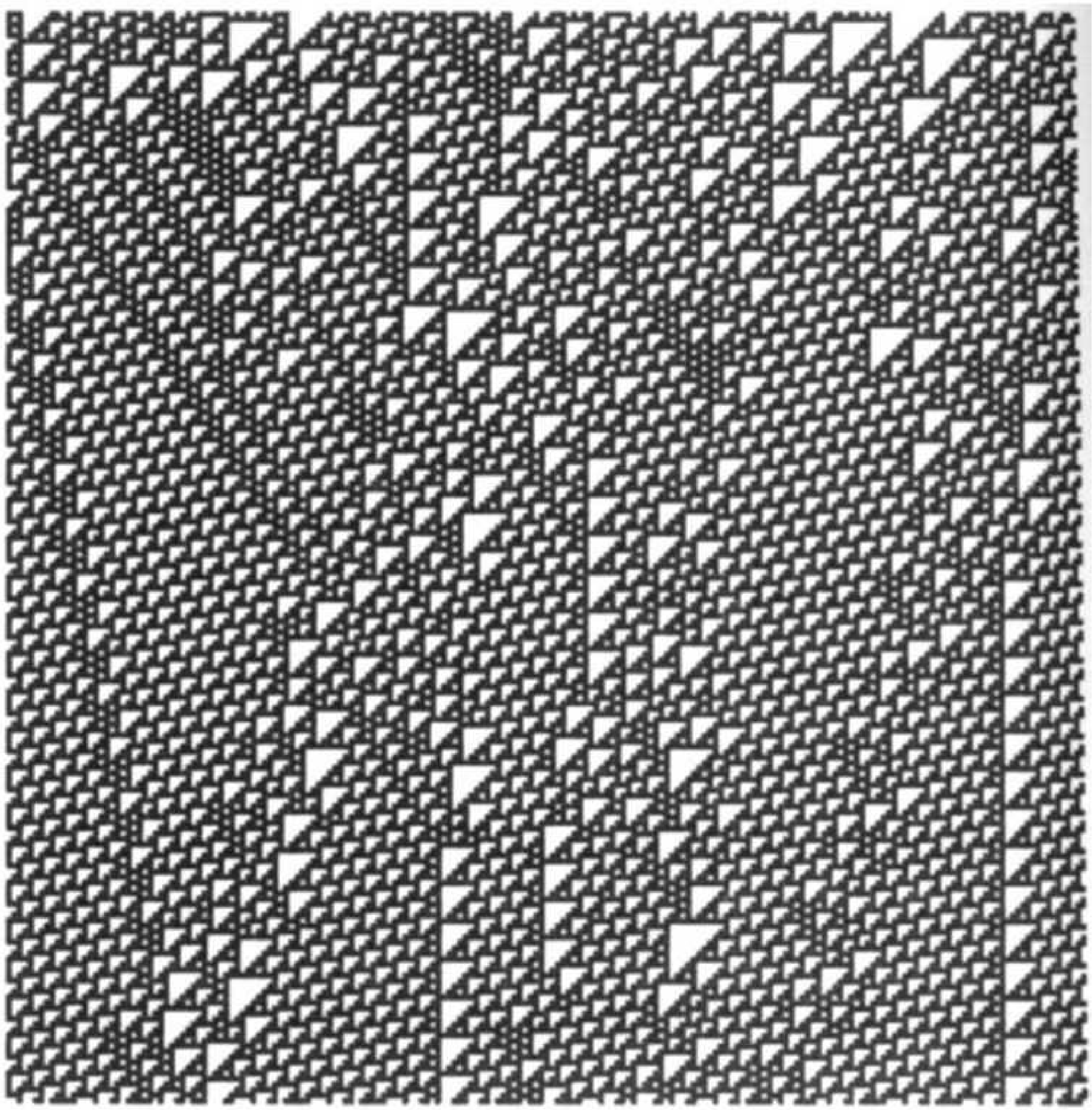


图10.7　规则110的时空图。这个一维元胞自动机有200个元胞，图中是从随机初始状态开始，200时间步的变化情况

看着图10.7和图10.8，也许你会理解为什么沃尔夫勒姆会对初等元胞自动机这么着迷。这种极为简单的元胞自动机规则究竟是如何产生出如此复杂的图样的呢？

对沃尔夫勒姆来说，简单的规则涌现出如此的复杂性，这简直就是神迹。他后来说："规则30自动机是我在科学中所遇到的最让人惊异的事物…… [152] 我花了几年时间来理解它的重要性。最后，我意识到这幅

图包含了所有科学长久以来的一个谜团的线索：自然界的复杂性到底从何而来。”沃尔夫勒姆对规则30印象非常深刻，[153] 以至于他用其来构造伪随机数发生器，并申请了专利。

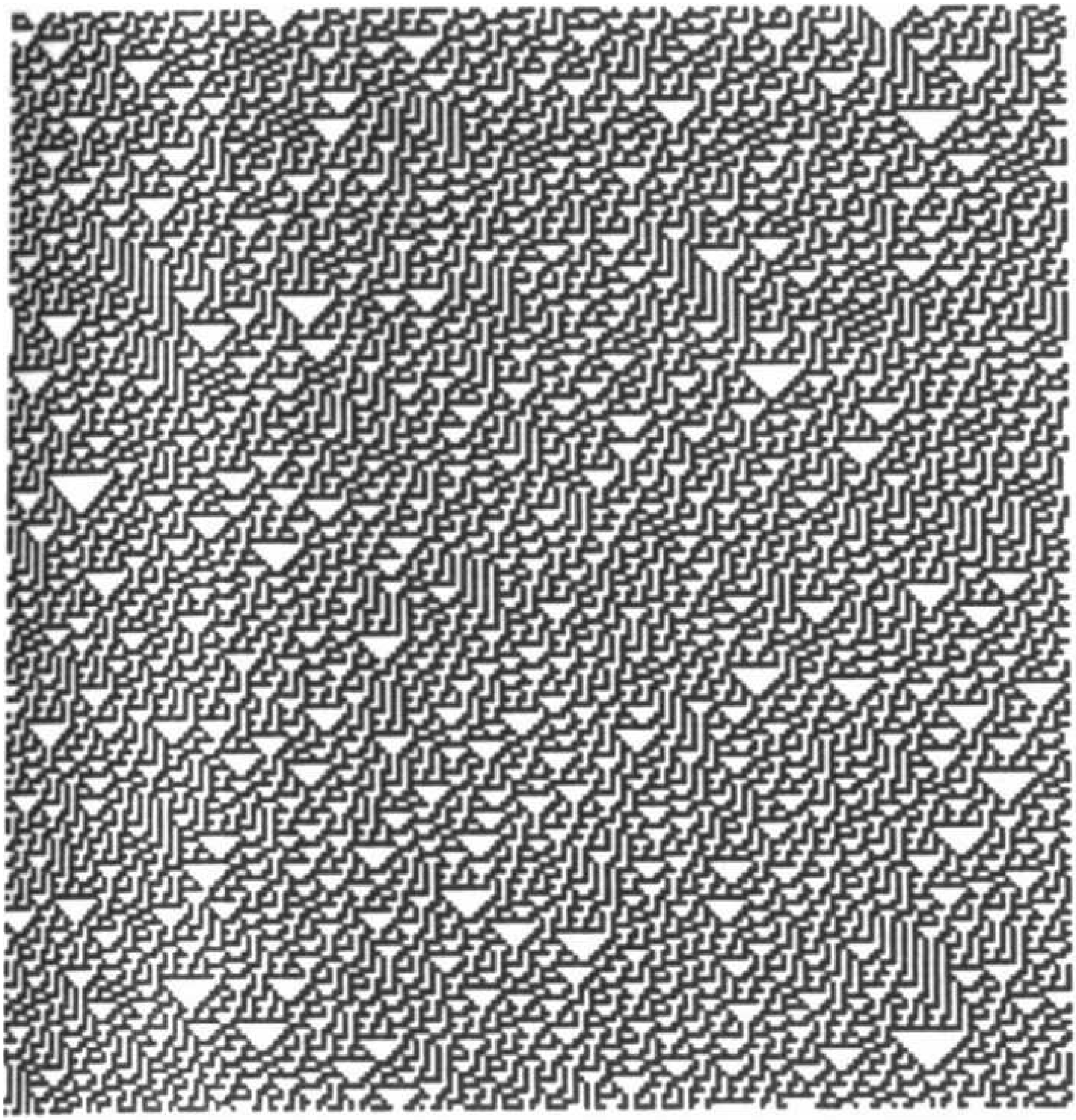


图10.8　规则30的时空图，从随机初始状态开始运行

沃尔夫勒姆对全部256种初等元胞自动机进行了彻底研究，从各种不同的初始状态开始，观察它们的变化。他让元胞自动机运行较长一段时间，直至元胞自动机的变化稳定下来。他发现最后都进入了4种类型

的变化情况：

类型1：不管初始状态如何，最后几乎都停止在不变的最终图样。规则8就是一个例子，不管初始状态如何，所有元胞很快就都变成了关状态，不再变化。

类型2：不管初始状态如何，最后要么停止在不变的图样，要么在几个图样之间循环。具体的最终图样依赖于初始状态。

类型3：大部分初始状态会产生看似随机的行为，虽然也会出现三角形等规则结构。规则30（图10.8）就属于这一类。

类型4：最有趣的类型。沃尔夫勒姆这样描述："类型4是有序与随机的混合：[154] 局部结构相当简单，但是这些结构会移动，并以非常复杂的方式相互作用。"规则110（图10.7）就属于这一类。

沃尔夫勒姆猜想，由于图样和相互作用的复杂性，类型4的所有规则都能进行通用计算。不过很难证明某个具体的元胞自动机、图灵机或其他机器是通用的。图灵证明的只是存在通用图灵机，冯·诺依曼也只是证明自复制自动机同时也是通用计算机。后来有几位科学家证明了简单的元胞自动机（比如生命游戏）是通用的。20世纪90年代，沃尔夫勒姆的一位研究助手库克（Matthew Cook）最终证明了规则110的确是通用的，[155] 这也许是目前已知的最简单的通用计算机。

沃尔夫勒姆的"新科学"

1998年，库克在圣塔菲研究所的一次会议上做报告，我第一次知道了他的结果。我当时的反应同我的大多数同事差不多，认为"太酷了！太巧妙了！不过没有什么实际或科学意义"。

和生命游戏一样，规则110也是极为简单的确定性系统产生出无法预测的复杂行为的例子。但在实际中很难设置一个初始状态来产生出所希望的复杂计算。而且规则110会比生命游戏更慢。

沃尔夫勒姆对这个结果的看法完全不同。在2002年出版的《一种新科学》（*A New Kind of Science* ）中，沃尔夫勒姆将规则110的通用性视为"新的自然定律" [156] ——他提出的计算等价性原理（Principle of

Computational Equivalence）——的有力证据。沃尔夫勒姆提出的这个原理包括4部分：

1.思考自然界中的过程的正确方法是将它们视为计算。

2.像规则110这样极为简单的规则（或“程序”）都能进行通用计算，这表明通用计算的能力在自然界中广泛存在。

3.通用计算是自然界中计算的复杂性的上限。也就是说，自然系统或过程不可能产生出“不可计算”的行为。

4.自然界中各种过程实现的计算在复杂程度上都几乎等价。

明白了没有？我必须承认，很难解释清楚这个原理的意思，沃尔夫勒姆这本1200页的鸿篇巨著，一个主要目的就是阐释这个原理，并说明它如何适用于各个科学领域。我通读了这本书，但还是没有完全明白沃尔夫勒姆的意思。不过我还是尽力解释一下。

沃尔夫勒姆说的“自然界中的过程就是计算”指的是类似于图10.7和图10.8中的某种东西。在任意给定时刻，元胞自动机都在通过将其规则应用于其当前状态来处理信息。沃尔夫勒姆认为，自然系统正是以这样的方式运作——它们包含信息，并根据简单规则处理这些信息。在《一种新科学》中，沃尔夫勒姆探讨了量子力学、进化和发育生物学、经济等领域，他想说明这些领域都能描述为使用简单规则进行的计算。本质上，他的“新科学”指的是这样的思想，宇宙和其中的万事万物都能用这种简单的程序来进行解释。这就是大写的计算，非常大。

因此，根据沃尔夫勒姆的观点，既然规则110这样极度简单的规则都能支持通用计算，那大部分自然系统——基本上应该都比规则110复杂——也就能支持通用计算。而且沃尔夫勒姆相信，给定正确的输入，没有比通用计算机所能进行的计算更复杂的计算。因此这就是自然界中所有可能计算的复杂性的上限。

第4章曾说过，图灵证明了通用计算机原则上能计算一切“可计算”的东西。但是一些计算要比其他的更简单。虽然都能在同样的计算机上运行，但“计算1+1”的程序肯定没有模拟地球气候的程序复杂，对吧？但沃尔夫勒姆的原理却说，实际进行的所有计算的“复杂程度”本质

上都是等价的。

这到底是什么意思呢？沃尔夫勒姆的理论还没有被广泛接受。这里我将给出我自己的意见。对于前两点，我认为沃尔夫勒姆提出的简单的计算机模型和实验能解释科学中许多过程的观点是正确的，这本书中的例子也证明了这一点。在第12章我还会讲到，我认为可以用信息处理解释许多自然系统的行为。我也发现许多这样的系统似乎的确能支持通用计算，虽然这一点的科学意义目前还不得而知。

至于第3点，目前也无法知道是否自然界中存在比通用计算机更强大的过程（能计算“不可计算的”东西）。现在已经证明，如果能建造出真正的非数字计算机（能计算具有无穷位小数的数字），你就能用其来解决停机问题 [157] ——我们在第4章看到的图灵不可计算问题。一些人，包括沃尔夫勒姆，不相信自然界真的存在无穷小数——也就是说，他们认为大自然本质上是数字的。两边都没有真正令人信服的证据。

第4点对我来说没有意义。我认为很有可能我的大脑支持通用计算（至少在我有限的记忆容量允许的范围内），而线虫的脑也（基本上）是通用的，但我不认为我们进行的计算在复杂程度上是等价的。

沃尔夫勒姆走得更远，他认为存在一个简单的类似元胞自动机的规则可以作为“宇宙的终极确定性模型” [158] ，这个原初元胞自动机的计算是存在的万事万物的源头。这个规则有多长呢？沃尔夫勒姆说：“我猜其实相当短。” [159] 但是到底多长呢？他说：“我不知道。也许三四行Mathematica程序。”小写的计算。

《新科学》在2002年出版后引起了轰动——很快就位居亚马逊网站的畅销书榜首，并在之后很久都留在畅销书榜中。沃尔夫勒姆为了这本书的出版成立了沃尔夫勒姆媒体公司（Wolfram Media），在书出版后这家公司进行了大规模的宣传活动。对这本书的看法分成两个极端：一些读者认为这本书棒极了，是革命性的；另一些人认为这本书盲目自大，缺乏实质和原创性[例如，批评者指出物理学家祖斯（Konrad Zuse）和弗瑞德金（Edward Fredkin）早在20世纪60年代就提出了宇宙是元胞自动机的理论 [160]]。不管怎样，我们这些元胞自动机的爱好者都认为，《新科学》至少很好地宣传了元胞自动机。

第11章　粒子计算[161]

1989年我偶然读到物理学家帕卡德（Norman　Packard）的一篇文章，[162] 写的是用遗传算法设计元胞自动机规则。我一下就被吸引住了，想自己也试试。当时因为繁忙，无暇顾及（我在写博士论文），但这个想法一直留在我的脑海里。几年后，论文完成了，我也在圣塔菲研究所找到了职位，终于可以有时间深入研究这个问题了。有个叫赫拉贝尔（Peter Hraber）的本科生当时在研究所逗留，想找点事情做，我就把他招进来协助我研究这个课题。很快一个叫达斯（Rajarshi Das）的研究生也加入我们，他当时在独立研究类似的课题。

类似于帕卡德的思想，我们想用遗传算法演化出能执行所谓的“多数分类[163]（majority classification）”任务的元胞自动机规则。这个任务很简单：元胞自动机要能区分初始状态中是开状态还是关状态占多数。如果是开状态占多数，最后所有元胞就应当都变成开状态。同样，如果是关状态占多数，最后所有元胞就应当都变成关状态。（如果初始状态中开状态和关状态的数量一样多，就没有答案，但是可以让元胞的数量为奇数来避免这种可能。）多数分类任务有点类似于选举，是在大家都只知道最近邻居的政治观点的情况下预测两个候选人谁会赢。

多数分类问题对冯·诺依曼结构的计算机而言是小菜一碟。CPU只需要分别对初始状态中的开状态和关状态进行计数，同时在内存中记录计数值就可以了。计数结束后，从内存中读取数值进行比较，然后根据结果将元胞状态都设成开或关。冯·诺依曼结构的计算机可以轻松实现这个工作，因为它有随机存取存储器可以存储初始状态和中间值，还有中央处理器可以计数，进行最后的比较，以及将状态重设。

而元胞自动机则没有CPU和内存可以用来计数。它只有一个一个的元胞，每个元胞除了自己的状态就只知道相邻元胞的状态。这种情形其实也是对许多实际系统的理想化。例如，在大脑中，神经元只与其他少数神经元有连接，而神经元必须决定是否激发，以及以何种强度激发，使得大量神经元的整体激发模式能表示特定的感知输入。类似的，第12章我们还会看到，蚂蚁必须根据与其他少量蚂蚁的交互来决定做什么事情，让蚁群整体能够受益。

因此，基本上很难设计出让所有元胞能协同决策的元胞自动机。赫

拉贝尔和我想知道遗传算法是不是能解决这个设计问题。

借鉴帕卡德的工作，我们使用一维元胞自动机，每个元胞与相邻的6个元胞相连，这样元胞的邻域中就有7个元胞（包括自己）。

你可以先想一下如何给元胞自动机设计规则，让它能进行多数分类。

一个合理的想法是："元胞应当变成邻域中当前占多数的状态。"这就好像根据你自己和邻居的多数意见来预测哪个候选人会当选。然而，这个"局部多数投票"元胞自动机并不能完成任务，图11.1说明了这一点。初始状态中黑色元胞占多数，然后根据局部多数投票规则运行了200步。可以看到元胞自动机很快变成黑白区域相间，然后就不再变化。黑白区域的边界两边的元胞邻域分别是黑白占多数。最后的状态组合中既有黑色元胞也有白色元胞，没有得到想要的结果。其中的问题是，根据这个规则，区域之间无法相互交流信息，因此它们无法得知自己是否是多数。

这个规则的设计并不是那么显而易见，因此依照第9章的做法，我们用遗传算法来试一下是不是能产生出有用的规则。

在遗传算法中元胞自动机规则被编码成0/1序列。每一位对应一种可能的邻域状态组合（图11.2）。

遗传算法的初始群体为随机产生的元胞自动机规则。为了计算规则的适应度，遗传算法用各种初始状态组合进行测试。遗传算法的适应度为产生出的最终状态正确的比例——正确指的是如果初始开状态占多数，元胞就都为开，初始关状态占多数，元胞就都为关（图11.3）。我们将遗传算法运行了许多代，最终算法设计出了一些表现得相当好的规则。

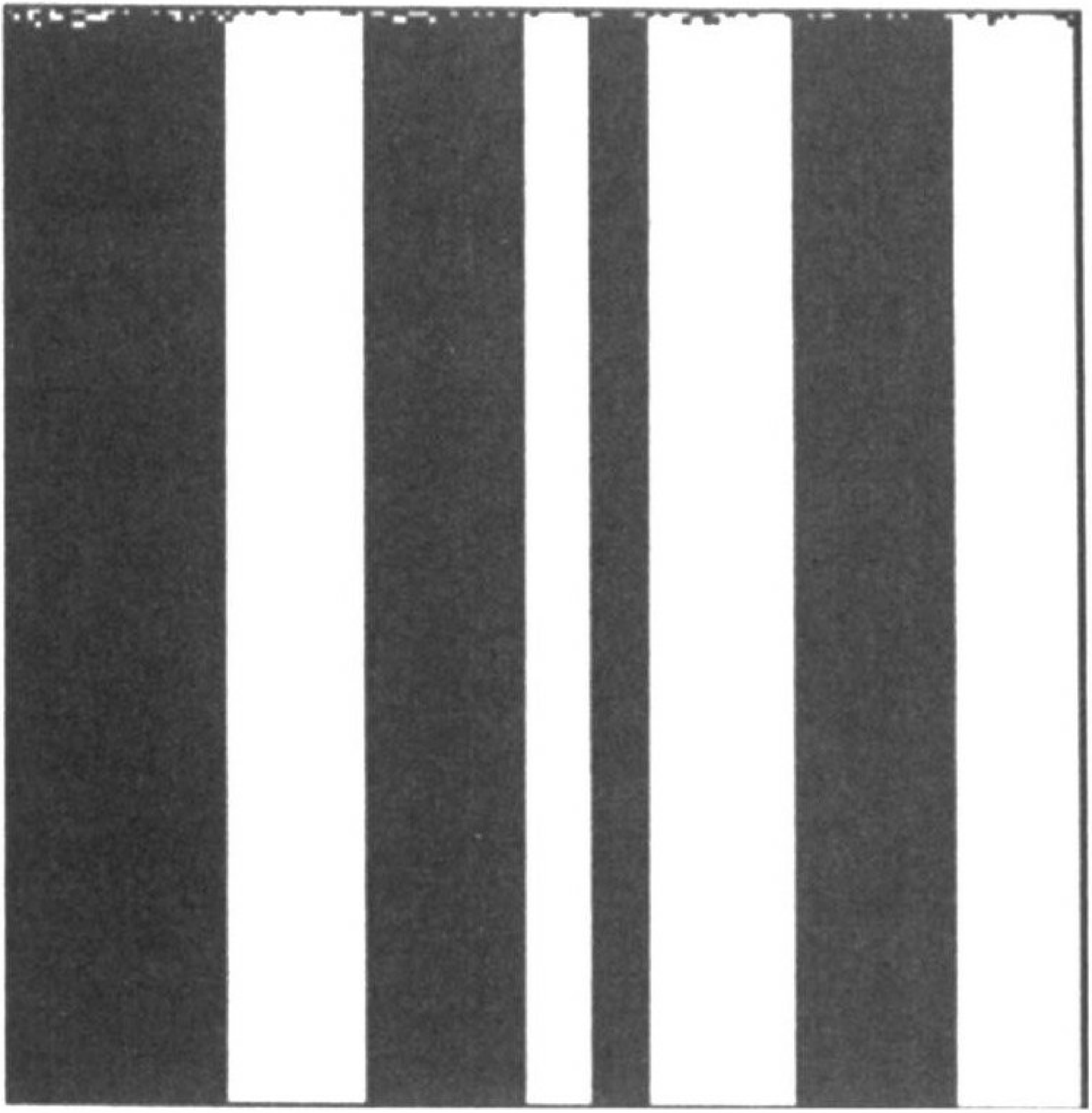


图11.1 “局部多数投票”元胞自动机的时空图，初始状态为黑色占多数（图片引自米歇尔、克鲁奇菲尔德和达斯的文章《演化执行计算的元胞自动机：最近的研究综述》，收录在《第一届进化计算及其应用国际会议文集》，俄罗斯科学院，1996年）

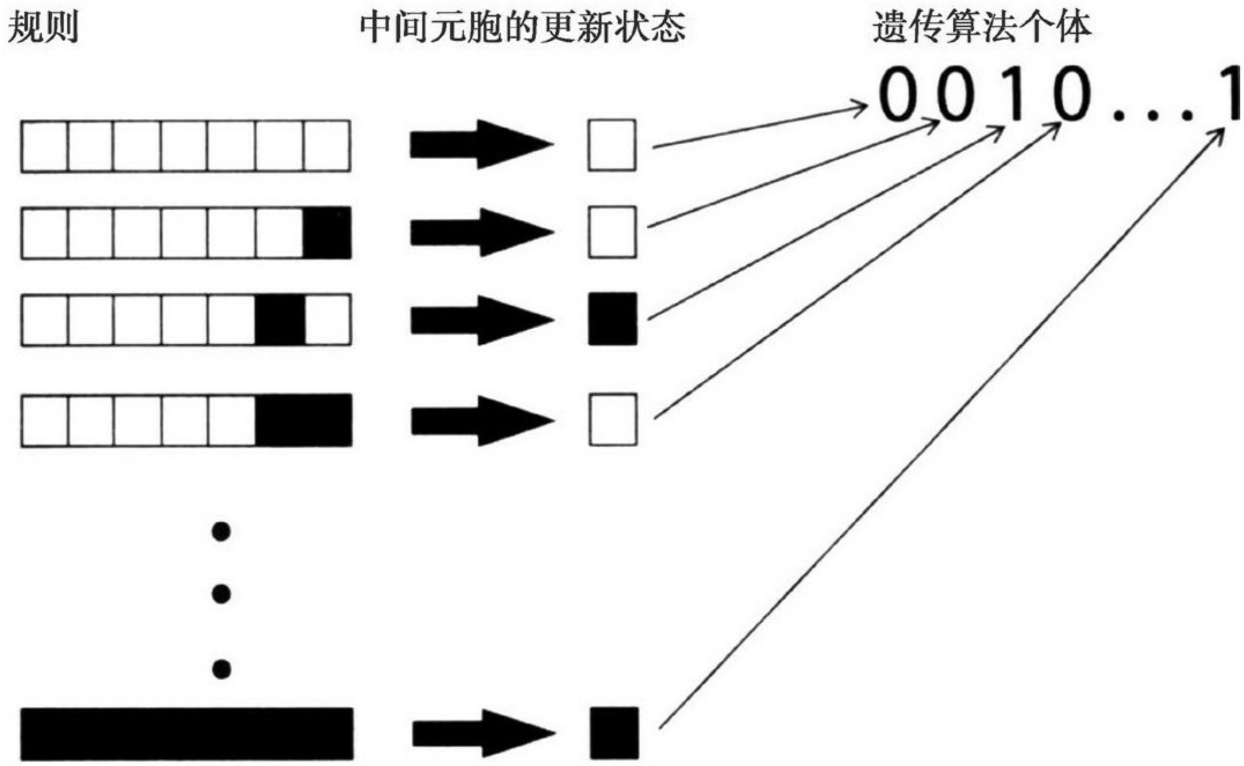


图11.2　元胞自动机编码为遗传算法个体的方法图示。128种可能的邻域状态组合按固定顺序排列。各邻域状态组合对应的中心元胞更新状态编码为0（关）和1（开）。遗传算法中每个个体有128位，以固定顺序编码更新状态

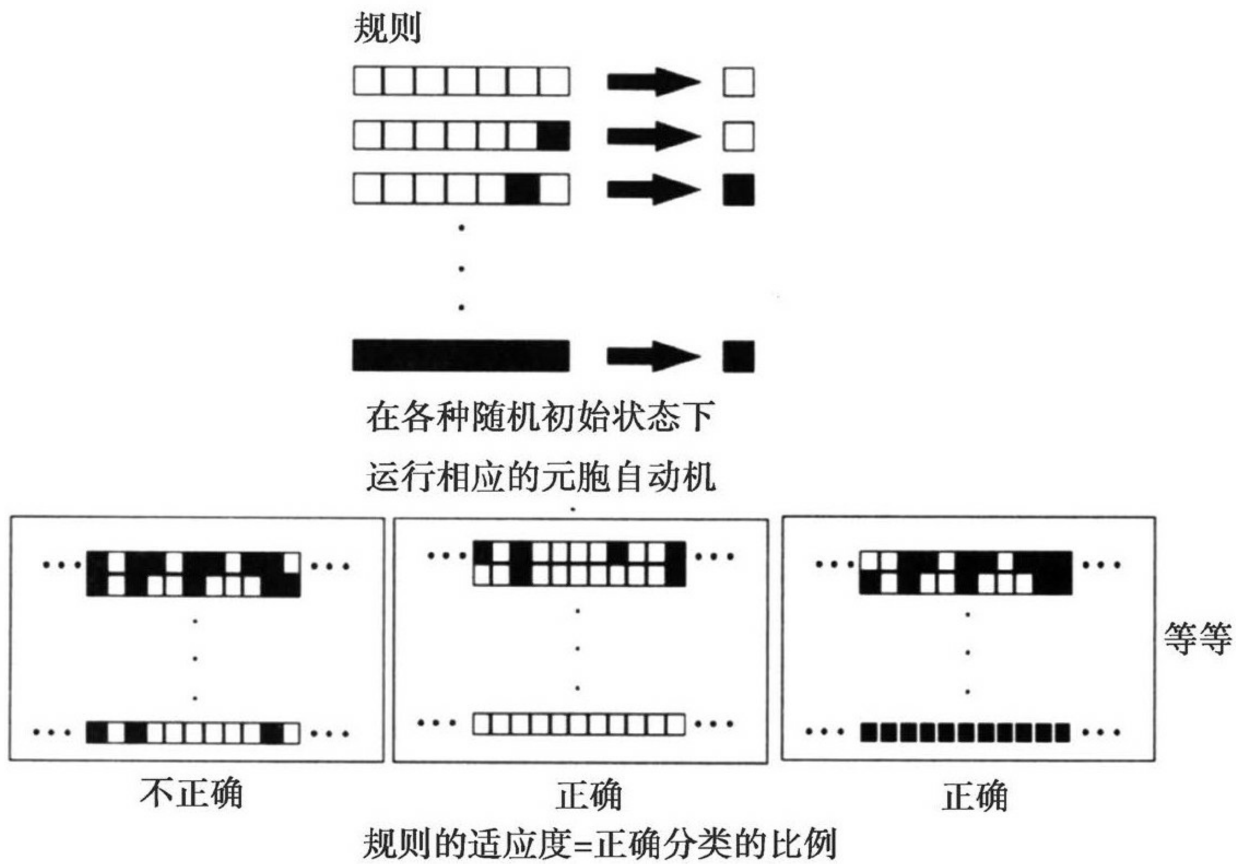


图11.3　为了计算适应度，用各种随机初始状态对规则进行测试。适应度为规则在运行一定步数后产生出正确答案的比例

我们通过机器人罗比已经看到，遗传算法演化得出的结果往往无法在其“基因组”层面进行理解。我们在这里演化出的用于多数分类的元胞自动机也是一样。下面是遗传算法演化出的表现最好的基因组之一：

0000010100000011000010101100001110000011100000010000010101010
1011101100100011101110000001010000000010111110111111111101101
1101111111

第1位是邻域全为0时中间元胞的更新状态，第2位是邻域为0000001时中间元胞的更新状态，依次往后。由于邻域状态有128种可能，因此基因组有128位。但光看这些数位是看不出这个规则如何运作，也无法知道为何它进行多数分类时适应度很高。

图11.4给出了这个规则在两种不同初始状态下的行为：分别为（a）黑色元胞占多数，（b）白色元胞占多数。可以看到在两种情形下最终

行为都是正确的——（a）变成了全黑，（b）变成了全白。在得到最终状态组合的过程中，元胞之间似乎在协同处理信息，以达到正确的最终状态。在这个过程中的图样很有意思，它们到底意味着什么呢？

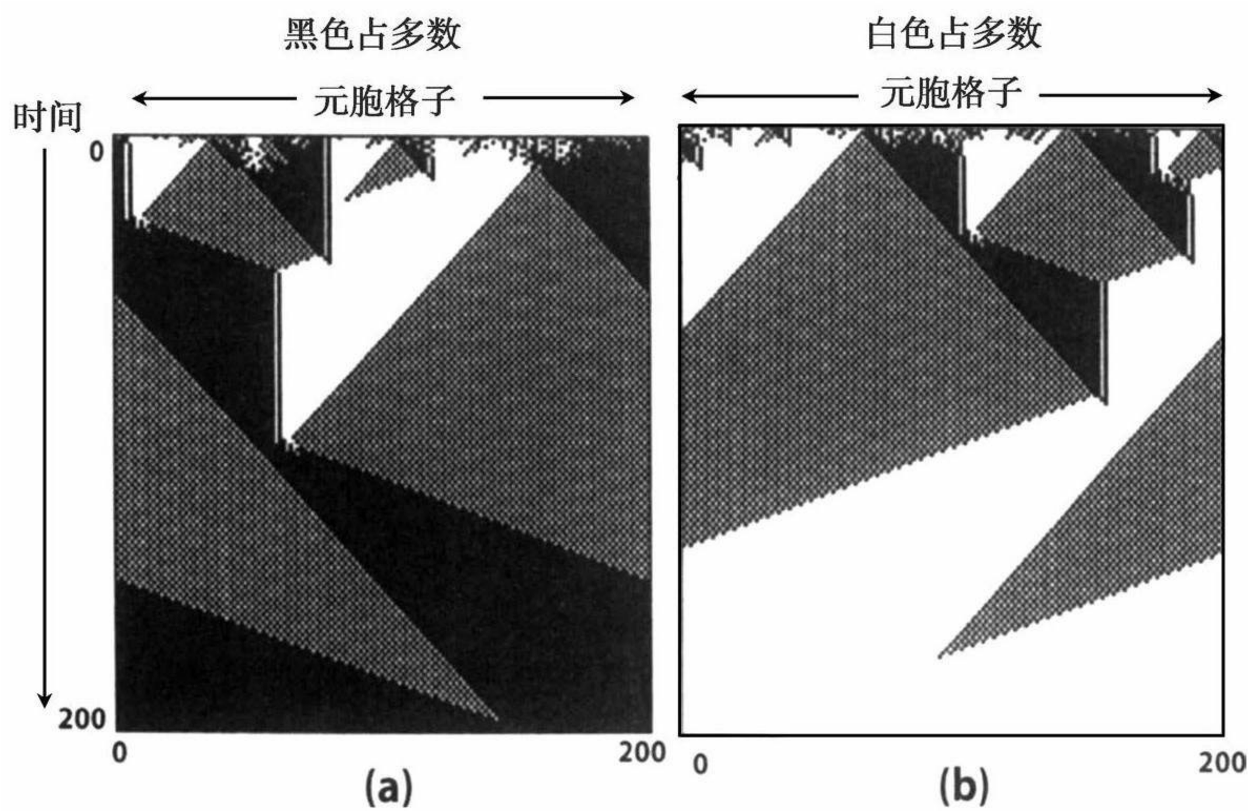


图11.4　演化出的表现最好的元胞自动机规则之一执行多数分类任务时的时空行为。在图（a）中，初始状态中黑色元胞占多数，元胞自动机迭代运行后全部变成了黑色。在图（b）中，初始状态中白色元胞占多数，元胞自动机迭代运行后全部变成了白色（图片引自米歇尔、克鲁奇菲尔德和达斯的文章《演化执行计算的元胞自动机：最近的研究综述》，收录在《第一届进化计算及其应用国际会议文集》，俄罗斯科学院，1996年）

我们花了很多时间盯着图11.4这样的图看，想知道到底发生了什么。幸运的是，伯克利的物理学家克鲁奇菲尔德碰巧访问了圣塔菲研究所，并对此产生了兴趣。他和他的同事正好发展了合适的概念能帮助我们理解这些图样是如何完成计算的。

图11.4中有三种类型的图样：全黑、全白以及类似于棋盘格的黑白交替（图11.4分辨率较低，所以看上去像灰色）。正是这种棋盘格图样传递了局部区域黑白元胞密度的信息。

同遗传算法为机器人罗比演化的策略一样，这个元胞自动机的策略

也相当聪明。图11.5是我对图11.4（a）做的标记。大部分为黑色和大部分为白色的区域很快就变成了全黑和全白。注意当黑色区域在左边而白色区域在右边时，中间的分界线总是垂直的。但如果白色区域在左边而黑色区域在右边，则会形成由黑白元胞相间组成的棋盘格图样三角形。如果把元胞自动机的两头回绕连在一起，就能看到三角形的环绕效果。

棋盘格区域的A边和B边以同样的速度扩展（即向两边覆盖同样的宽度）。A线向西南延伸，直到碰到垂直界线。B线则刚好避开了从另一边与这条垂直界线相撞。这表明A线的延伸长度要短一些，也就是说，A线左边的白色区域要小于B线右边的黑色区域。碰撞后，A线消失了，而黑色区域则得以继续扩展。在三角形底部的角上，B线和C线消失了，全部元胞格子都变成了黑色，从而得到了正确的状态组合。

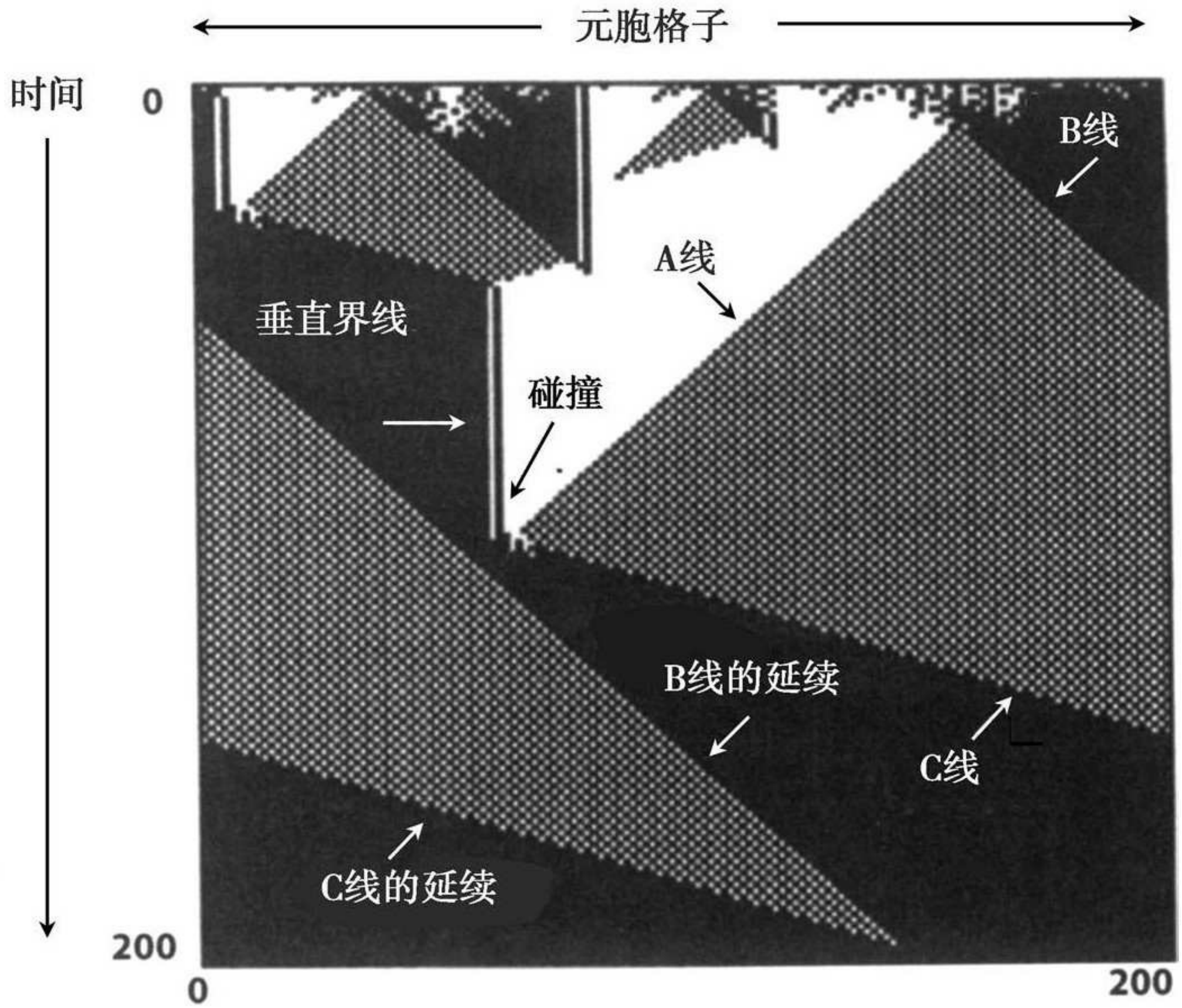


图11.5　标记了重要特征后的图11.4（a）

如果我们将这些图样当作是进行计算，那么垂直界线和棋盘格区域就可以视为信号。这些信号通过元胞的局部互动而产生和传播。信号使得元胞自动机能作为一个整体判断相邻的大片黑白区域哪个更大，截去较小的区域，并使较大的区域得以扩展。

这些信号就是图11.1中的局部多数投票元胞自动机与图11.5中的元胞自动机的主要区别。前面提到，前者的黑色和白色区域之间没有信息沟通，因而也无法得知两者谁占多数。而对于后者，棋盘格和垂直边界产生的信号则可以进行这样的信息传递，通过信号之间的相互作用对传递的信息进行处理，从而得出答案。

克鲁奇菲尔德之前发明了一种方法 [164] 可以研究动力系统行为中

的“信息处理结构”，他建议我们用这种方法来分析遗传算法演化出的元胞自动机。克鲁奇菲尔德的思想是区域之间的界线（例如图11.5中的A、B、C线以及垂直界线）携带有信息，而界线发生碰撞就是对信息进行处理。图11.6对图11.5的时空图进行了重绘，将黑色、白色和棋盘格区域都去掉，只留下界线，这样更容易看清楚一些。这张图有点像以前物理实验用的气泡室中的基本粒子轨迹。因此克鲁奇菲尔德就把这些界线称为“粒子”。

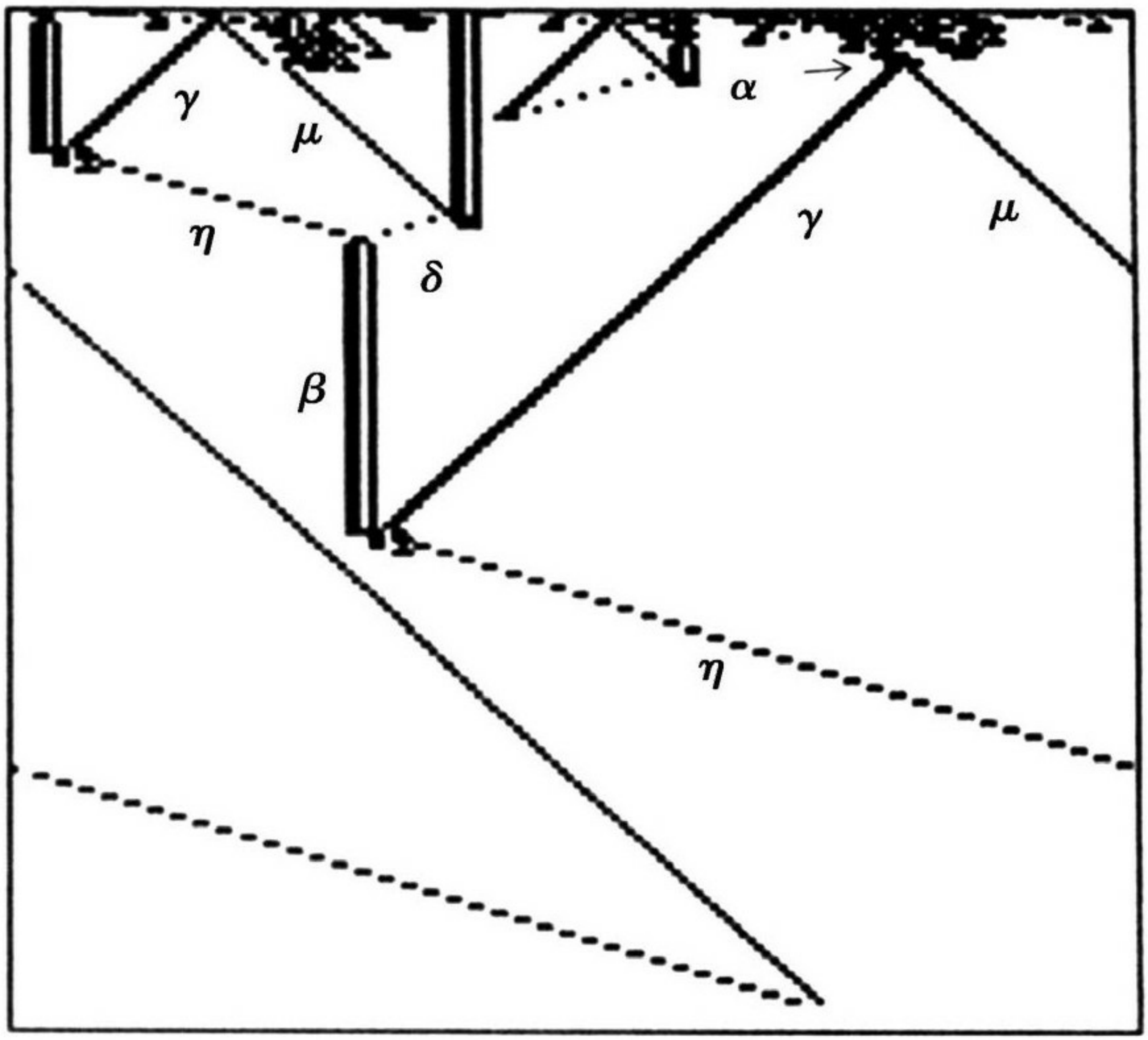


图11.6　对时空图图11.4（a）进行重绘，将单一的图样去掉，只留下区域之间的界线（“粒子”）（图片引自米歇尔、克鲁奇菲尔德和达斯的文章《演化执行计算的元胞自动机：最近的研究综述》，收录在《第一届进化计算及其应用国际会议文集》，俄罗斯科学院，1996年）

物理粒子通常习惯用希腊字母表示，这里也照样。这个元胞自动机

能产生6种不同类型的粒子：γ（伽马）、μ（缪）、η（艾塔）、δ（德尔塔）、β（贝塔）、α（阿尔法，一种寿命很短的粒子，衰变成γ和μ）。每种粒子分别对应一种界线——例如，η与黑色和棋盘格区域之间的界线相对应。粒子碰撞有5种，其中3种（β+γ、μ+β、η+δ）会产生出新粒子，而另外两种（η+μ和γ+δ）则会“湮灭”，两个粒子都消失。

将元胞自动机的行为用粒子进行描述能帮助我们理解其如何编码信息和进行计算。例如，α和β粒子编码了关于初始状态的不同类型的信息。α粒子衰变成γ和μ。γ粒子携带了白色区域边界的信息；类似的，μ粒子则携带了黑色区域边界的信息。当γ抢在μ之前与β碰撞，β和γ所携带的信息就结合到了一起，从而可以得知，最初由这几条界线划分的大块白色区域要小于大块黑色区域。这个新的信息被编码在新产生的粒子η中，它的任务就是追上粒子μ，并与其一起湮灭。

这种分析方法也可以被应用于其他几种演化出来的执行多数分类等任务的元胞自动机。这使得我们能预测特定元胞自动机的适应度等特征（而不用去运行元胞自动机本身，只需要分析其“粒子”描述就可以了）。另外我们也能理解为何某个元胞自动机的适应度会较高，以及如何描述各种元胞自动机在执行计算时所犯的错误。

粒子描述让我们看到了仅仅观察元胞自动机规则或是其时空图变化看不到的东西：它们让我们能从信息处理的角度来解释元胞自动机是如何执行计算的。粒子是我们强加给元胞自动机的描述，而不是在元胞自动机中确实发生的事情，遗传算法也没有用它们来演化元胞自动机。然而不知怎的，遗传算法却演化出了行为可以用粒子信息处理解释的规则。事实上粒子以及它们之间的相互作用可以作为一种语言，用来解释以一维元胞自动机为背景的分布式计算。沃尔夫勒姆曾提出《元胞自动机理论的二十个问题》 [165] ，其中最后一个问题是：“对元胞自动机的信息处理能不能给出高层次的描述？”他想要的可能就是类似于这种语言的东西。

这些都是最近的工作，还需要继续深入研究。我相信，用这种方法理解计算，虽然不符合正统，却对没有中央控制，分布在简单个体中的计算会有用。例如，关于感知信号的高级信息在大脑中如何编码和处理目前仍然是个谜。也许可以用某种类似于粒子的语言对其进行解释，也有可能是类似于波的计算，因为大脑是三维的，神经元在一起形成携带

信息的波运动，并通过波的相互作用处理信息。

大脑计算当然与一维元胞自动机不在一个层面上。不过，有一种自然系统却可以用非常类似于粒子的语言解释：植物的气孔网络。所有有叶植物的叶子表面都布满了气孔——根据光线和湿度开合的微小孔隙。气孔打开时可以让二氧化碳进来，用于光合作用。但是气孔打开也会导致植物体内的水分蒸发。犹他州立大学（Utah State University）的植物学家莫特（Keith Mott）、物理学家皮克（David Peak）和他们的同事长期观察叶子气孔的开合模式[166]，他们认为气孔组成了一个有点类似于二维元胞自动机的网络。他们还发现气孔开合的时间模式很像二维形式的粒子相互作用。他们猜测植物通过气孔进行分布式计算——通过优化气孔的开合让二氧化碳的获取和水分流失达到最佳平衡——这种计算也许也能用粒子语言进行解释。

第12章　生命系统中的信息处理[167]

自从西拉德发现信息能将热力学第二定律从麦克斯韦妖的威胁下拯救出来之后，信息和计算就日渐成为科学的宠儿。在许多人看来，信息具有本体地位，同质量和能量一样，被当作实在的第三种基本成分。在生物学中尤其如此，将生命系统描述成信息处理网络已成为潮流。信息处理一词随处可见，以至于你可能会认为它的意义已没有疑义，也许就是基于香农对信息的形式化定义。然而，同其他复杂系统科学的核心概念一样，信息处理的概念也很不清晰；一旦脱离了图灵机和冯·诺依曼结构计算机的精确形式化背景，就经常很难厘清信息处理或计算的概念。上一章介绍的工作就是在元胞自动机的背景下针对这个问题的尝试。

这一章的目的是探讨生命系统中的信息处理或计算。我将描述三种不同的自然系统，免疫系统、蚁群和细胞代谢——在其中信息处理似乎都扮演了关键的角色——并尽力阐明信息和计算在它们中所扮演的角色。最后我将尝试弄清在此类分散系统中信息处理的共性。

什么是信息处理

我先引用一下自己在第10章的一段话：“自然系统的‘计算’指的是什

么呢？大致上说，计算是复杂系统为了成功适应环境而对信息进行的处理。但是这样的说法还能更精确些吗？信息在哪里？复杂系统又是如何处理信息的？”这些问题似乎很显然，但如果深究一下，我们很快就会陷入复杂系统科学最深的泥沼之中。

当我们说一个系统在处理信息或计算（从现在开始，我会不加区分地使用这两个词）时，我们就面临着以下问题：[168]

◆“信息”在这个系统中扮演了什么角色？

◆信息又是如何传递和处理的？

◆这些信息是如何获得意义的？又是对谁有意义？（有些人会不同意计算需要某种类型的意义，不过我会先坚持己见，认为它需要。）

传统计算机中的信息处理

在第4章我们已经知道，20世纪30年代图灵用图灵机对输入的处理步骤定义了计算的概念。图灵的定义是传统冯·诺依曼结构计算机的设计基础。对于这些计算机，关于信息的问题很容易回答。我们可以说信息就是带子上的符号和读写头的可能状态。信息的处理则是通过读写头在带子上的读写和状态变化实现的。这一切都是根据规则进行的，程序就是由规则组成的。

对于传统计算机的程序，我们（至少）可以从两个层面来看：机器码层面和编程语言层面。在机器码层面上，程序是由具体的让机器一步一步执行的低级指令组成（例如，“将内存中地址n的数据移到CPU的寄存器j”，“对CPU寄存器j和i中的数据执行或逻辑运算，将结果存入内存中地址m处”，等等）。而在编程语言层面上，程序是由BASIC或JAVA这样的高级语言的指令组成，让人更容易理解（例如，“将某个变量乘以2，并将结果赋给另一个变量”，等等）。一个高级语言指令通常要用几个低级指令来实现，不同的计算机类型可能有不同的实现。因此高级语言程序可以以不同的方式实现为机器码；高级语言是对信息处理更抽象的描述。

图灵机输入和输出的信息的意义来自于人们（程序员或使用者）的解读。中间步骤产生的信息的意义也来自人们对高级语言命令步骤的解

读（或设计）。高级语言描述让我们能容易理解在机器码或硬件层面上对于人来说很抽象的计算。

元胞自动机中的信息处理

对于元胞自动机等非冯·诺依曼结构的计算机来说，答案就不是那么显而易见了。例如我们在上一章用遗传算法演化出来执行多数分类任务的元胞自动机就是这样。与传统计算机做个类比，我们可以说元胞自动机的信息就是元胞格子在每一步的状态组合。输入就是初始状态组合，输出则是最终的状态组合，在每个中间步的信息则根据元胞自动机规则在元胞邻域内进行传递和处理。意义来自人们对所执行的任务的认识以及对从输入到输出的映射的解读（例如，“元胞最终都变成了白色；这意味着初始状态组合中白色元胞占多数”）。

在这个层面上描述信息处理就类似于在“机器码层面”进行描述，并不能帮助人们理解计算是如何完成的。同冯·诺依曼结构计算的情形一样，在这里我们也需要一种高级语言来理解中间步骤的计算，对元胞自动机底层的具体细节进行抽象。

上一章我提出，粒子以及粒子的相互作用可以用来描述元胞自动机的信息处理，类似于高级语言。信息通过粒子的运动来传递，粒子的碰撞则是对信息进行处理。这样，信息处理的中间步骤就通过人们对粒子行为的解释获得了意义。

冯·诺依曼结构的计算之所以容易描述，一个原因就是，编程语言层面和机器码层面可以毫无歧义地相互转化，因为计算机的设计让这种转化可以很容易做到。计算机科学提供了自动编译和反编译的工具，让我们可以理解具体的程序是如何处理信息的。

而元胞自动机则不存在这样的编译和反编译工具，至少目前还没有，也没有实用和通用的设计“程序”的高级语言。用粒子来帮助理解元胞自动机高级信息处理结构的思想也是最近才出现，还远没有形成此类系统的计算理论体系。

理解元胞自动机信息处理的困难在实际生命系统中也同样存在。“自然系统的‘计算’指的是什么？”对于这个问题目前仍然知之甚

少，在科学家、工程师和哲学家之间存在着广泛的争议。然而它对于复杂系统科学又是极为重要的问题，因为对生命系统中信息处理的高层次描述不仅能让我们从更高的视角理解具体系统的运作，也能让我们超越系统繁杂的细节，抽象出一般性原理。实质上，这种描述就是为生物学提供一种“高级语言”。

这一章余下部分将力图用具体的例子来阐明这种思想。

免疫系统

在第1章曾讲到过免疫系统。现在我们来更深入地了解一下，为了保护身体免受病毒、细菌、寄生虫等病原体的侵害，免疫系统[169]是如何处理信息的。

我们知道，免疫系统是由数以亿计的各种细胞和分子组成，它们在身体里循环，通过各种信号相互影响。

在免疫系统各种类型的细胞中，我关注的是淋巴细胞（白细胞的一种，图12.1）。淋巴细胞是在骨髓中产生的。有两种淋巴细胞很重要，释放抗体攻击病毒和细菌的B细胞，以及杀死入侵者同时还调节其他细胞反应的T细胞。

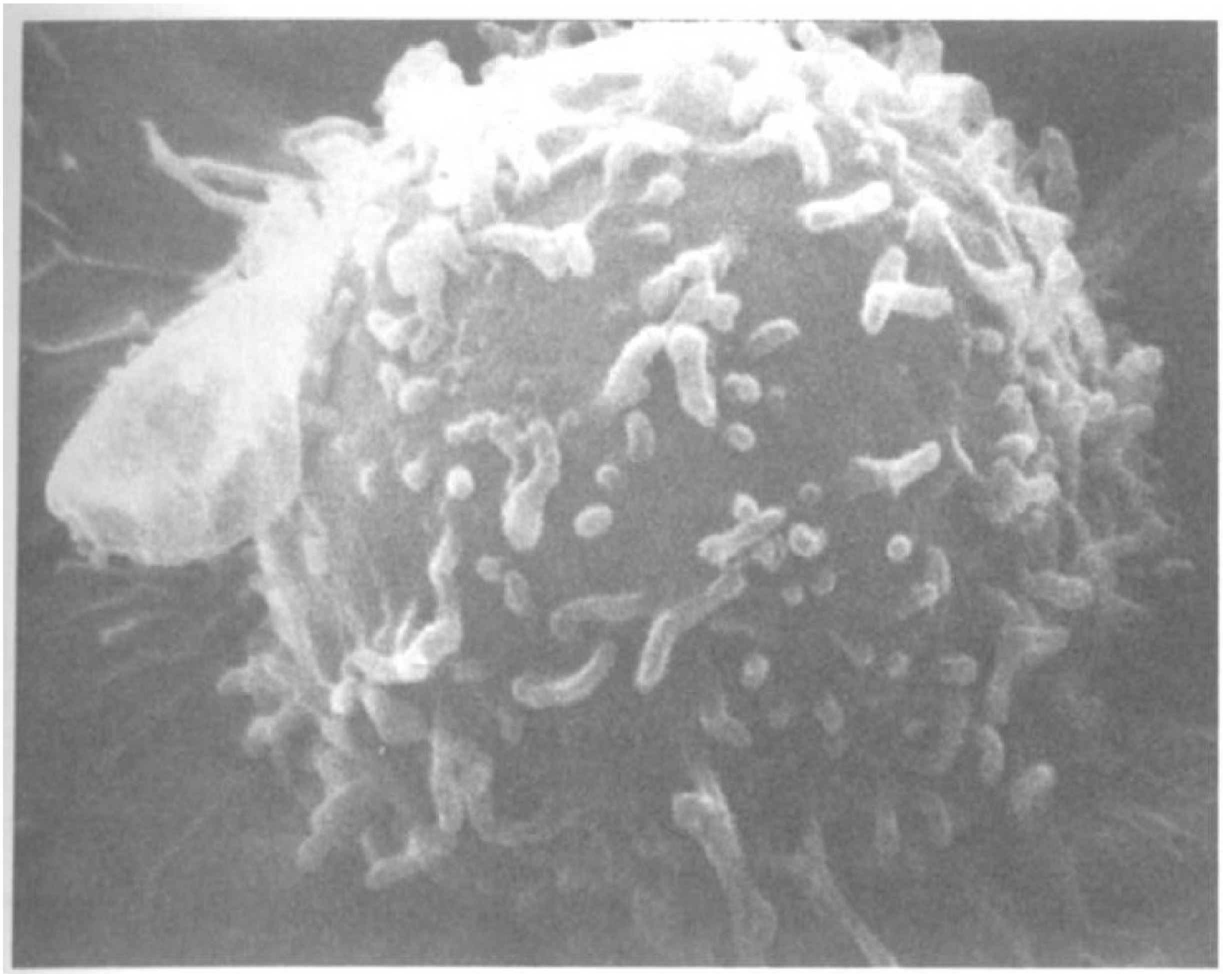


图12.1　人类淋巴细胞，表面覆盖着受体，可以与遇到的特定形状的分子相结合[图片来自美国国家癌症研究所（National Cancer Institute, http：//visualsonline.cancer.gov/details.cfm?imageid=1944）]

身体中所有细胞表面都有称为受体的分子。顾名思义，这些分子是细胞接收信息的途径。信息表现为能与受体分子结合的外界分子。受体能否与某个分子结合取决于它们的分子结构是否能充分匹配。

一个淋巴细胞表面覆盖的受体是一样的，可以与特定的某一类分子形状匹配。如果恰好遇到了形状相匹配的病原体分子（称为“抗原”），淋巴细胞的受体就会与其相结合，淋巴细胞就“识别”出了抗原，这是消灭病原体的第一步。结合可强可弱，依赖于分子与受体的匹配程度。图12.2描绘了这个过程。

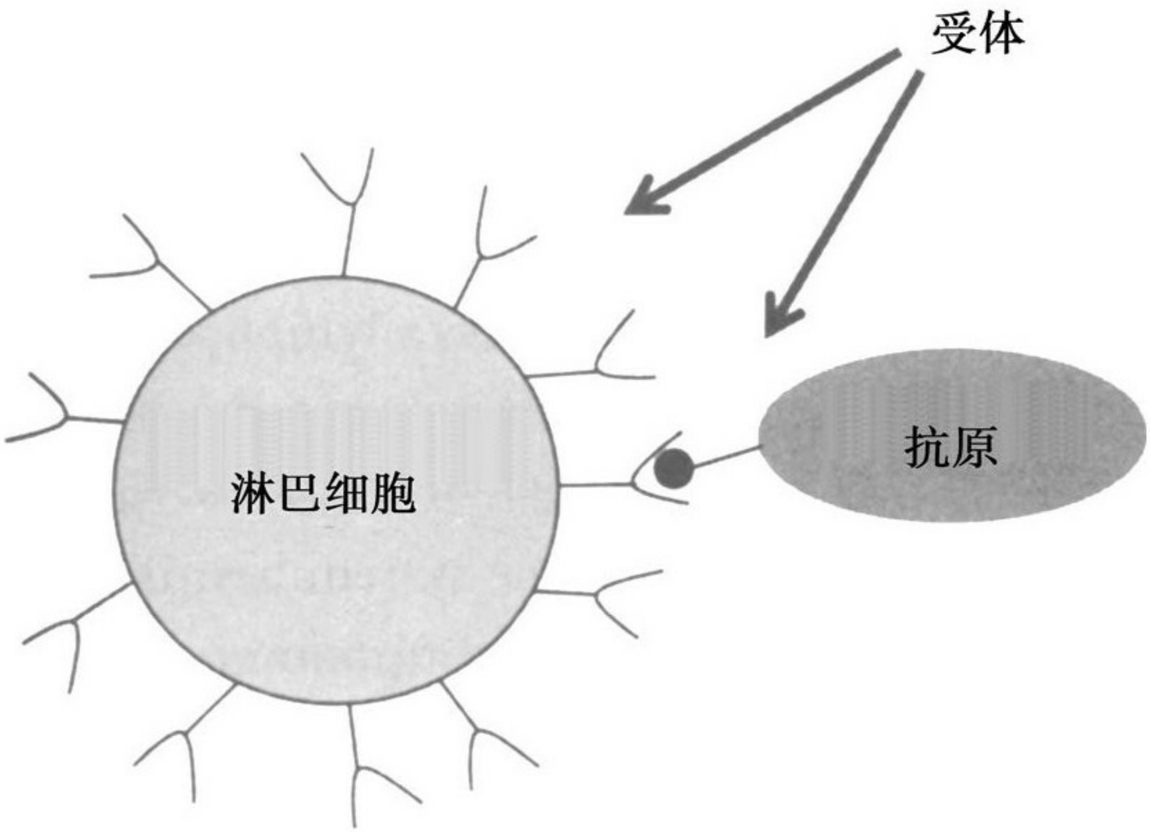


图12.2　淋巴细胞（图中为B细胞）受体与抗原结合的示意图

免疫系统面临的主要问题是，它不知道什么病原体将会入侵身体，因此它也就不可能“预先设计”出一组淋巴细胞，让它们的受体与入侵病原体的分子形状刚好能紧密结合。而且可能的病原体种类是个天文数字，因此免疫系统永远也不可能在同一时间产生出那么多淋巴细胞以应对每一种可能。虽然身体每天会产生数以百万计不同的淋巴细胞，但是身体可能遇到的病原体却还要多得多。

我们来看看免疫系统是如何解决这个问题的。为了能“覆盖到”各种各样可能的病原体外形，身体内会同时存在许多种类型的淋巴细胞。免疫系统利用随机性，让淋巴细胞能识别的形状范围互不相同。

当淋巴细胞产生时，通过淋巴细胞DNA复杂的随机重组过程，新的受体会被创造出来。由于淋巴细胞群体不断更新（每天会产生上千万新的淋巴细胞），身体也就不断产生具有新的受体形状的淋巴细胞。对于任何进入体内的病原体，身体很快就产生出能与病原体的标记分子（也就是抗原）相结合的淋巴细胞，虽然结合可能不是很紧密。

一旦发生了结合事件，免疫系统就得搞清楚这是不是真正的威胁。病原体当然是有害的，一旦它们进入身体，就会开始大量复制。不过发动免疫系统攻击会导致发炎等对身体有害的症状，攻击太强烈甚至有可能致命。免疫系统作为一个整体必须确定威胁是否足够严重，值得承担让免疫反应伤害身体的风险。免疫系统只有在强结合事件足够多之后才会进入高速运转模式。

B细胞和T细胞这两种类型的淋巴细胞协同工作，判断攻击是否有必要。一旦B细胞表面强结合受体的数量超过了某个阈值，与此同时B细胞从有类似受体的T细胞那里收到了“发动”信号，B细胞就会被激活，表明它现在感觉到了身体受到威胁（图12.3）。一旦激活，B细胞就会向血液中释放抗体分子。这些抗体与抗原结合，使它们失效，并对它们进行标记，好让其他免疫细胞摧毁它们。

激活的B细胞被输送到淋巴结，在那里迅速分裂，产生出大量后代，复制时由于变异，许多后代的受体形状都改变了。然后这些后代会与淋巴结俘获的抗原进行测试。不能结合的细胞很快就会死去。

存活下来的后代被释放到血液中，其中一些会遇到抗原并与其结合，有时候会比它们的母细胞结合得更紧密。这些激活的B细胞同样又被输送到淋巴结，在那里产生出自己的后代。这也就是为什么当你病了的时候淋巴结会肿大，因为它在大量产生淋巴细胞。

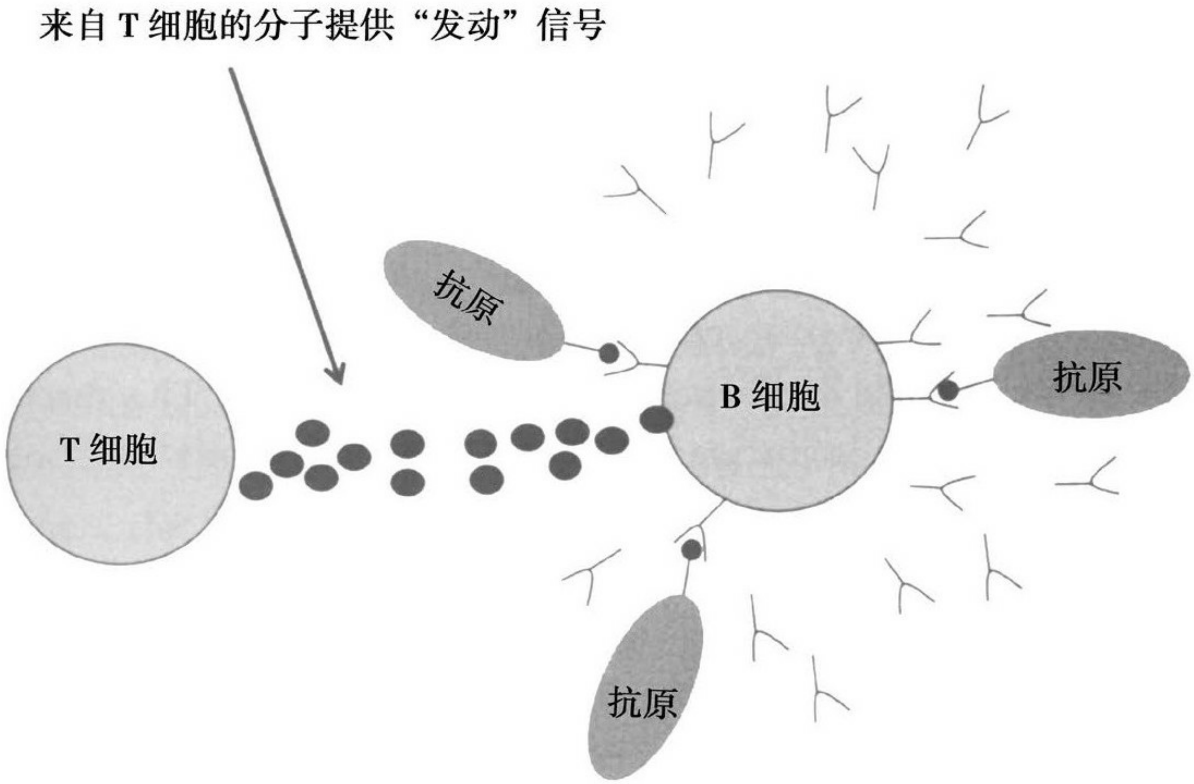


图12.3　通过与抗原结合和接收T细胞的“发动”信号激活B细胞的示意图。信号刺激B细胞产生和释放抗体（y形分子）

这个循环不断进行，与抗原匹配得越好的B细胞产生的后代也越多。简而言之，这就是一个自然选择过程，B细胞群体进化出能与目标抗原紧密结合的受体形状，从而使得通过选择“设计”出来攻击特定抗原的抗体武器库不断扩大。

这个侦测和摧毁过程一般需要数天或数周时间才能将目标病原体从身体中清除干净。

这种策略起码存在两个潜在问题。首先，免疫系统如何防止淋巴细胞错误攻击身体自身的细胞？其次，免疫系统在对身体的伤害作用太大时如何停止或调整攻击？

免疫学家们还没有完全弄清这些问题，现在这些问题都还是活跃的研究领域。有人认为避免攻击自身的一个主要机制是所谓的负选择（negative selection）。当淋巴细胞产生出来时，它们不会被立即释放到血液中去，它们会在骨髓和胸腺中进行测试，与身体自身的分子进行接

触。与“自身”分子紧密结合的淋巴细胞可能会被杀死或对基因进行“编辑”以改变受体。也就是说免疫系统只使用不会攻击自身的淋巴细胞。这个机制经常会失效，有时候会产生出糖尿病或类风湿性关节炎这类自身免疫性疾病。

另一个防止自身免疫攻击的主要机制可能是调节性T细胞（regulatory T cells）的作用。调节性T细胞是T细胞的一个特殊亚种。目前还不清楚调节性T细胞是如何工作的，只知道它们通过某种化学机制抑制其他T细胞的活动。还有一种可能的机制是B细胞之间对一种有限资源的竞争[170]——一种称为B细胞刺激因子（BAFF）的特殊化学物质，B细胞需要它才能存活。在负选择过程中漏网并且仍然与自身分子紧密结合的B细胞，由于总是与自身分子结合在一起，因此比其他B细胞需要更多的BAFF。对这种有限资源的竞争导致结合自身分子的B细胞死亡的可能性增加。

虽然免疫系统攻击外来病原体，但它也还是有义务在攻击的毒性和尽可能防止伤害身体之间进行平衡。免疫系统使用了一系列机制来实现这种平衡（目前对这些机制还知之甚少）。其中许多机制都依赖于一组信号分子，被称为细胞因子（cytokines）。对身体的伤害会导致细胞因子的分泌，细胞因子会抑制活跃的淋巴细胞。可能伤害越严重，细胞因子的浓度就越高，活跃的细胞也就越有可能遇到它们，从而被关闭，达到调节免疫系统的目的，而不用对整个免疫系统进行抑制。

蚁群

第1章曾说过，蚁群与大脑很相似。都可以看作由相对简单的个体（神经元、蚂蚁）组成的网络，并且涌现出宏观尺度上的信息处理行为。有两个例子表现出蚁群的这种行为，一是以最佳和适应性的方法搜寻食物的能力，以及适应性地根据蚁群整体的需要分配蚂蚁执行各种工作。这两种行为都是在没有中央控制的情况下完成的，所使用的机制与前面描述的免疫系统惊人的相似。

大部分蚂蚁种类的食物搜索大致是这样进行的。[171]蚁群中搜寻食物的蚂蚁随机朝一个方向搜索，如果遇到食物，就返回蚁穴，沿途留下作为信号的化学物质——信息素（pheromones）。当其他蚂蚁发现了信息素时，就有可能会沿着信息素的轨迹前进。信息素的浓度越高，蚂蚁

就越有可能跟着信息素走。如果蚂蚁找到了那堆食物，就会返回巢穴，将信息素的轨迹增强。如果信息素的轨迹得不到增强，就会消失。通过这种方式，蚂蚁一起创造和沟通关于食物位置和质量的各种信息，并且这种信息还会适应环境的变化。存在的轨迹和强度很好地表达了搜索蚁协同发现的食物情报（图12.4）。



图12.4　蚁迹（Flagstaffotos版权所有，经许可重印）

蚁群的任务分配也是以分散方式进行的。生态学家戈登（Deborah Gordon）曾研究过红色收割蚁（Red Harvester ants）的任务分配。[172]蚁群中的工蚁分为四个工种：搜寻食物、维护蚁穴、巡逻和垃圾处理。执行各种任务的工蚁数量能随着环境变化。戈登发现，如果蚁穴被稍微搅乱，维护蚁穴的工蚁数量会增加。如果附近的食物源很多，质量很好，搜寻食物的工蚁数量就会增加。单只蚂蚁可以根据蚁穴环境的变化做出适应性响应，决定采取哪种工作，无需另外的蚂蚁来指挥，每只蚂蚁也仅与其他少数蚂蚁交互，它们是如何做到的呢？

答案可能是蚂蚁根据它们周围的环境以及它所遇到的执行各种任务的蚂蚁比例来决定自己干什么。比如，一只闲逛的蚂蚁——目前什么也没有做——在蚁穴附近遇到了杂物，它执行蚁穴维护工作的概率就会增加。另外，如果它发现很多维护蚁穴的工蚁在进进出出，也会增加执行蚁穴维护工作的概率；因为这种活动的增加表明有重要的蚁穴维护工作在进行。类似的，维护蚁穴的工蚁如果遇到了很多搜寻食物的蚂蚁带着种子返回蚁穴，就会增加它转向搜寻食物工作的概率；因为种子搬运信号的增加表明发现了高质量的食物源，需要进行采集。显然，通过用触须与其他蚂蚁交流，侦测与各项工作有关的特殊化学物质，蚂蚁就能知道其他蚂蚁在做什么。

类似的这种利用信息素与其他个体直接交互的机制可能也是其他种类蚂蚁和社会昆虫集体行为的基础，例如第1章中看到的蚂蚁用身体搭桥和构建庇护所，[173]这些行为的许多方面还有待进一步研究。

生物代谢

新陈代谢是指一系列化学过程，生物消耗从食物、空气或阳光中获取的能量，维持生命所需的所有功能。这些化学过程大部分发生在细胞内部，通过称为代谢途径的化学反应链进行。在生物体内的每个细胞中，营养分子通过反应产生能量，细胞组分也通过代谢途径产生。这些组分维持和修复内部的以及外部的功能和细胞间通信。数以百万的分子不断在细胞质中随机运动，分子不断相互碰撞，偶尔（微秒级尺度）酶会碰到形状匹配的分子，从而加速由酶控制的化学反应。这样逐级反应逐渐形成大分子。

淋巴细胞通过释放细胞因子影响免疫系统，蚂蚁通过释放信息素影

响收集食物的行为，同样，代谢途径上的化学反应也会不断改变特定途径的速度和获得的原材料。

代谢途径一般是复杂的化学反应序列，受自我调节反馈控制。例如，糖酵解（glycolysis）是所有生命中都存在的代谢途径——它通过多步反应将葡萄糖转化为丙酮酸（pryruvate），丙酮酸又通过称为柠檬酸循环的代谢途径产生许多物质，其中包括ATP（三磷腺苷），ATP是细胞能量的主要来源。

这种代谢途径的数量数以百计，有些独立，有些相互依赖。代谢途径生成新的分子，开启其他代谢途径，并调节自身或其他代谢途径。

与前面描述的免疫系统和蚁群的调节机制类似，代谢调节机制也是基于反馈。糖酵解就是这样的例子。糖酵解的一个主要作用是为制造ATP提供必需的化学原料，如果细胞中ATP的量很多，就会减缓糖酵解的速度，从而降低ATP的产生速度。反过来，如果细胞缺乏ATP，糖酵解的速度就会加快。代谢途径的速度一般都受途径产生的化学物质调节。

这些系统中的信息处理

现在我们来尝试回答这一章开始时提出的关于信息处理的问题：

◆“信息”在这些系统中扮演了什么角色？

◆信息是如何被传递和处理的？

◆信息是如何获得意义的？又是对谁有意义？

信息扮演了什么角色

在元胞自动机的例子中，当我说到信息处理时，我指的不是细胞、蚂蚁或酶这样的单个个体的行为，而是一大群这种个体的集体行为。根据这个框架，信息不像在传统计算机中那样，位于系统中的某个具体位置。在这里它表现为系统组分的动态模式和统计结果。

在免疫系统中，淋巴细胞的空间分布和时间动力学可以解释为体内

病原体数量变化信息的动态表示。类似的，细胞因子浓度的空间分布和动态可以解释为免疫系统杀死病原体和避免伤害身体的宏观尺度信息。

对于蚁群，食物源的信息则动态地表示为蚂蚁在不同蚁迹上的统计分布。蚁群的整体状态表示为执行各种任务的蚂蚁的动态分布。

至于细胞代谢，当前状态以及细胞需求的信息则不断通过各种分子的浓度和动态变化反映出来。

信息是如何被传递和处理的

通过采样实现通信

将信息编码为基本组分的统计和变化模式的一个后果是，没有哪个个体组分能感知或传达系统状态的“宏观画面”。信息必须通过空间和时间采样来传递。

在免疫系统中，淋巴细胞通过受体接受抗原和免疫系统其他细胞释放的细胞因子来对环境采样。淋巴细胞采样这些分子信号的空间和时间分布，进而被激活或是休眠。其他细胞反过来又会对激活的淋巴细胞的浓度和类型进行采样，并受其影响，将对抗病原体的细胞吸引到身体的特定区域。

对于蚁群，单只蚂蚁则是通过感受器对信息素信号进行采样。它们根据对环境中信息素浓度特征的采样决定移动方向。前面讲过，单只蚂蚁也会用基于浓度的采样信息——遇到的其他蚂蚁——来决定是否进行某项工作。在细胞代谢中，通过酶与特定分子结合，酶对分子浓度的时空变化进行采样，进而又反馈到代谢途径。

行为的随机成分

由于获得的信息具有统计性，系统组分的行为就必然是随机的（至少“不可预测”）。前面描述的三个系统本质上都利用了随机性和或然性。[174] 每个淋巴细胞的受体形状都有随机生成成分，从而能采样许多可能的形状。淋巴细胞随血流分布，因此体内淋巴细胞的空间分布也有随机成分，从而可以采样抗原的多种可能空间特征。淋巴细胞激活的具体阈值、实际的分裂速度以及后代的变异都具有随机性。

类似的，蚂蚁搜寻食物的活动也具有随机成分，蚂蚁对信息素的侦测以及是否受其轨迹吸引也是随机的。蚂蚁改变工种也是以随机的方式。生物化学家齐夫（Edward Ziff）和科学史学家罗森菲尔德（Israel Rosenfield）这样描述随机性的作用：“最终蚂蚁会建立通往食物源路径的详细地图。[175] 观察者可能会认为蚂蚁的食物分布地图是由某位智能设计者提供的。通往食物源的路径看上去就像是精心画出来的，但实际上却是一系列随机搜索的产物。”

细胞代谢依赖于分子的随机扩散和分子相遇的概率，随着系统的变化，相对浓度会发生变化，从而概率也会跟着变化。

为了让数量相对较少的简单个体（蚂蚁、细胞、分子）能探测相比起来要大得多的可能范围，这种内在的随机性和或然性似乎是必需的，尤其是通过探测获得的信息本质上是统计性的，而且对于将要遇到的事物也没有什么先验知识。

但是随机性必须与确定性达成平衡：复杂适应系统的自我调节不断调整各项事务的概率——个体应该向哪里移动，它们应当采取什么行动，以及如何探测庞大空间中的具体路径。

微粒化探测

复杂生物系统绝大多数都有微粒化结构，它们由大量相对比较简单的个体组成，个体以高度并行的方式协同工作。

这种结构有几个可能的好处，稳健、效率高、可以演化。还有一个额外的好处就是微粒化并行系统能进行侯世达所说的“并行级差扫描[176]（parallel terraced scan）”。他指的是对许多可能性和路径同时进行探测，某项探测所能获得的资源依赖于其当时的成效。搜索是并行的，许多可能性被同时探测，但是存在“级差”，意思是并不是所有可能都以同样的速度和深度进行探测。利用获得的信息不断调整探测，从而有所侧重。

例如，免疫系统需要随时确定，在病原体可能形状的庞大空间中，哪部分区域需要用淋巴细胞进行探测。体内每个淋巴细胞都可以看作形状空间的微小探测器。探测器的形状范围越成功（与抗原结合得越紧密），获得的探测资源也就越多，即后代淋巴细胞越多；形状范围不成

功的（结合不紧密的淋巴细胞）则得不到那么多资源。不过，在处理获得的信息的同时，免疫系统也在不断产生新的淋巴细胞，用来探测全新的形状范围。这样系统在应对当前状况的同时，也不会忽略新的可能性。

与此类似，蚂蚁搜寻食物时也使用并行级差扫描策略：开始时许多蚂蚁随机寻找食物。一旦在某个方向发现了食物，就会分派更多的系统资源（蚂蚁），通过前面描述的反馈机制，进一步探测这个方向。路径得到的探测资源不断通过其相对绩效——所发现食物的数量和质量——进行动态调整。但是，由于蚂蚁数量很多，再加上具有随机性，绩效不好的路径也会继续探测，当然分派的资源会少得多。谁知道呢，说不定就能发现更好的食物源。

在细胞代谢机制中，微粒化探测是由代谢途径实现的，每条途径执行特定的任务。代谢途径的反应速度受其本身和其他途径的反馈影响。反馈通过分子浓度变化的形式体现，从而使得各途径的相对速度可以不断根据细胞的当前需求进行调整。

此外，系统微粒化的特性不仅使其能探测各种不同的路径，同时也使得系统能够连续地调整探测路径，因为采取的动作都相对较小。而如果更加粗粒化，就很有可能在没有绩效的探测路径上浪费时间。因此，探测的微粒化特性使得系统能根据其获得的信息连贯地对探测进行调整。不仅如此，微粒化系统天生具有冗余度，因此即使有个体组分不能可靠工作，获取的信息也只是统计性的，系统还是能正常运转。冗余度使得对信息有许多独立的采样，而且只有大量组分采取同样的微粒化行动时才会产生效果。

分散探测与集中行动之间的互动

在三个例子中，根据系统的需要，随机分散探测与集中行动之间不断进行互动。

在免疫系统中，分散探测是通过带有各种受体、尝试匹配可能抗原的淋巴细胞群体的不断变化进行。集中行动则是让成功匹配的淋巴细胞产生后代，集中应对特定形状的抗原。

蚂蚁搜寻食物时是由蚂蚁随机移动、四处寻找食物来进行分散探

测，在集中行动中则是蚂蚁循着信息素轨迹活动。

在细胞代谢中，作为分散行动的分子随机探测则与由化学浓度和基因调节控制的集中激活或抑制结合在一起。

对于所有的适应性系统，在两种探测模式中保持适当的平衡都是关键[177]。而最优的平衡点随时间不断变化。开始时所知的信息很少，探测基本是随机分散的。随着信息增多并产生影响，探测逐渐变得具有确定性，集中于对系统的感知进行响应。简而言之，系统既要探测信息，又要对信息加以利用，不断调整适应。在分散探测和集中行动之间进行平衡可能是适应性和智能系统的共性。例如霍兰德就曾用这种平衡解释遗传算法的工作原理。

信息是如何获得意义的

信息如何获得意义（有些人称为目的性），这是哲学的一个永恒话题。对于哲学家们的看法我无从置评，但是要想理解生命系统中的信息处理过程，我们必须面对这个问题。

在我看来，意义与生存和自然选择密切相关。如果事件影响到某个生物的生存或繁衍能力，那么事件对生物就具有某种意义。总之，事件的意义是如何应对事件的依据。事件对生物免疫系统的意义是其对生物适应度的影响。（我在这里非正式地使用适应度一词。）这些事件对免疫系统有意义是因为它们告诉免疫系统该如何应对，以提高生物的适应度——对于蚁群、细胞以及其他生物的信息处理系统也是一样。聚焦于适应度是我理解意义的概念，并将其应用到生物信息处理系统的一条途径。

但是在前面描述的复杂系统中，并不存在中央控制或领导者，那么是谁或是什么在觉察当前情势的意义[178]，然后据此做出适当的反应呢？这个问题实际上问的就是什么构成了生命系统的意识或自我意识。对我来说，这个问题是复杂系统研究和整个科学最深的谜团。这个谜团是许多科学和哲学书的主题，但是至今还没有让人完全满意的答案。

将生命系统视为在进行计算的观点有个有趣的副产品：它激发了计算机学家编写程序模仿这类系统来完成真实任务。例如，免疫系统的信

息处理思想引出了所谓的人工免疫系统[179]：保护计算机免受病毒和各种入侵者攻击的适应性程序。类似的，蚁群启发了所谓的“蚁群算法[180]”，模拟蚂蚁释放信息素和转换工作的原理来解决移动电话路由优化和货运调度优化等困难问题。下一章我还将介绍我与我的博士导师合作研究的人工智能程序，这个工作是受包括细胞代谢在内的全部三个系统的启发。

第13章　如何进行类比（如果你是计算机）[181]

容易的事很难

有一天，我对8岁的儿子杰克说：“杰克，把袜子穿上。”他把袜子顶到头上，“你看，我把袜子穿上了！”他觉得很好玩。而我则意识到他的搞怪行为说明了人类和计算机之间一个很大的区别。

“把袜子穿到头上”的玩笑之所以好笑（至少对8岁的孩子来说是这样），是因为它违反了我们都知道的常识：人类的大部分言辞原则上讲都有些模棱两可，但是当你和别人说话时，他们还是知道你的意思。如果我对我丈夫说：“亲爱的，你知道我的钥匙在哪里吗？”如果他仅仅回答说“知道”，我会很恼火——显然我的意思是“告诉我，我的钥匙在哪里”。当我最好的朋友说她感到在工作中寸步难行时，我回应说“心有同感”，她会知道我的意思不是说我觉得她的工作寸步难行，而是说我自己的工作。这种相互理解就是所谓的“常识”，说得更正式点是“对上下文敏感”。

而现代计算机则对上下文一点也不敏感。我的计算机有一个最新的垃圾邮件过滤器，有时候却区分不出带有V！a&®@这样的“单词”的邮件有可能是垃圾邮件。最近《纽约时报》还报道，记者现在都在学习如何根据搜索引擎的特点而不是从读者的角度来设计标题：“一年前，《萨克拉门托蜂报》[182]（*Sacramento Bee*）改变了网上的版面标题。‘房地产’换成了‘商品房’，‘景致’换成了‘生活格调’，餐饮信息在纸版上是‘美食’，在网络版上则是‘美食/餐饮’。”

当然，并不是说计算机干什么都很蠢。在一些特定的领域它们已经变得相当聪明。计算机控制的汽车现在能自己穿越崎岖的沙漠。计算机

程序在对一些疾病的诊断上能胜过医生，在解复杂的方程时能胜过数学家，下棋能打败象棋大师。人工智能（AI）领域这种让人振奋的例子数不胜数。计算机学家霍维茨（Eric Horvitz）说：“在会议上你会听到人们说‘人类级的AI’，[183] 他们说起来毫不脸红。”

呵呵，也许是达到某些人的水平。不过还是有些事情计算机做不了，比如理解人类的语言，描述一张照片的内容，或是像前面讲的使用常识。明斯基（Marvin Minsky）是人工智能的先驱之一，他曾简明扼要地总结AI的悖论：“容易的事很难 [184] 。”计算机能做许多人类认为需要很高智商的事情，同时它们却又做不了三岁小孩都能做的事情。

进行类比

有一个很重要的能力现在的计算机还不具备，那就是进行类比。

说到类比，人们经常会回想起让人头痛的标准化考题，如：“鞋子对于脚就像手套对于___？”不过我说的类比范围更加宽泛：类比是在两个表面上不同的事物之间发现抽象的相似性的能力。这个能力渗透到了智能的几乎所有方面。

我们来看看下面的例子：

小孩能知道在图画书和照片上的小狗与真实的小狗都是同一个概念的实例。

人可以轻松辨认出各种印刷体和手写的字母A。

吉恩对西蒙妮说：“我每星期给爸妈打一次电话。”西蒙妮回应说“我也是这样。”显然她的意思不是她每星期给吉恩的父母打电话，而是说给自己的父母打电话。

一位女士对她的男同事说：“我最近的工作太忙了，都没有时间陪我丈夫了。”他回应说，“我也是一样”——他的意思不是说他忙得没时间陪这位女士的丈夫，而是说他没时间陪自己的女友。

一则广告将毕雷矿泉水（Perrier）说成是“瓶装水中的卡迪拉克”。一篇新闻将教学说成是“职业中的贝鲁特”，伊拉克则被说成

是“又一个越南”。

英国与阿根廷之间发生了福克兰岛（马岛）战争，福克兰岛靠近阿根廷，居民则太多数是英裔。希腊支持英国，因为它自己同土耳其对塞浦路斯岛存在争端，塞浦路斯靠近土耳其，居民则主要是希腊裔。

古典音乐爱好者在收音机上听到一段没听过的旋律，马上就能知道是巴赫的作品。早期音乐发烧友听到一段巴洛克管弦乐，很容易就能辨别出是哪个国家的作曲家。顾客能知道超市的背景乐改编自甲壳虫乐队的《嗨，朱迪》。

物理学家汤川秀树（Hideki Yukawa）用电磁力作类比来解释核力，并据此推断出对于核力也存在媒介粒子，其性质类似于光子。后来确实发现了这种粒子，性质也与预测相符。汤川秀树因此获得了诺贝尔奖。

在日常生活中和那种可遇不可求的发现中存在大量类比，这里只是一小部分例子。这些例子表明，人们在各种层面上都能很好地认识到两种事物和情形之间的类似之处，让各种概念从一种情形流畅地“滑到”另一种情形。这些例子揭示了人类思想这种独一无二的能力，就像19世纪哲学家梭罗（Henry David Thoreau）说的：“所有对真理的认识都是通过类比得来的[185]。”

计算机的类比能力可以说是臭名昭著。这也就是为什么我不能给计算机看一张小狗游泳的图片，然后要它在图片库中将“类似的照片”找出来。

我对类比的认识经历

20世纪80年代初，我大学刚刚毕业，还不是很明确今后该做些什么，我到纽约的一所中学当数学老师。这份工作的薪水很低，而纽约物价昂贵，因此我削减了不必要的开支。但我还是买了一本新出的书，作者是印第安纳大学（Indiana University）的一位计算机系教授，[186]题目有点奇怪，《哥德尔、艾舍尔、巴赫——集异璧之大成》。因为我的专业是数学，又参观过许多博物馆，所以知道哥德尔和艾舍尔是谁，而且我也喜欢古典音乐，所以对巴赫也很熟悉。但是将他们的名字放到一

起作为书的标题就让我搞不懂了，这勾起了我的好奇心。

没想到侯世达的书改变了我的一生。从题目看不出来，书的内容是思维和意识是如何从大量简单神经元的分散行为中涌现出来的，这类似于细胞、蚁群和免疫系统的涌现行为。这本书让我第一次了解了复杂系统的一些主要思想。

侯世达（图13.1）想用类似的原理建造有智能和“自我意识”的计算机程序。这很快也成了让我充满激情的目标，我决定去跟侯世达研究人工智能。

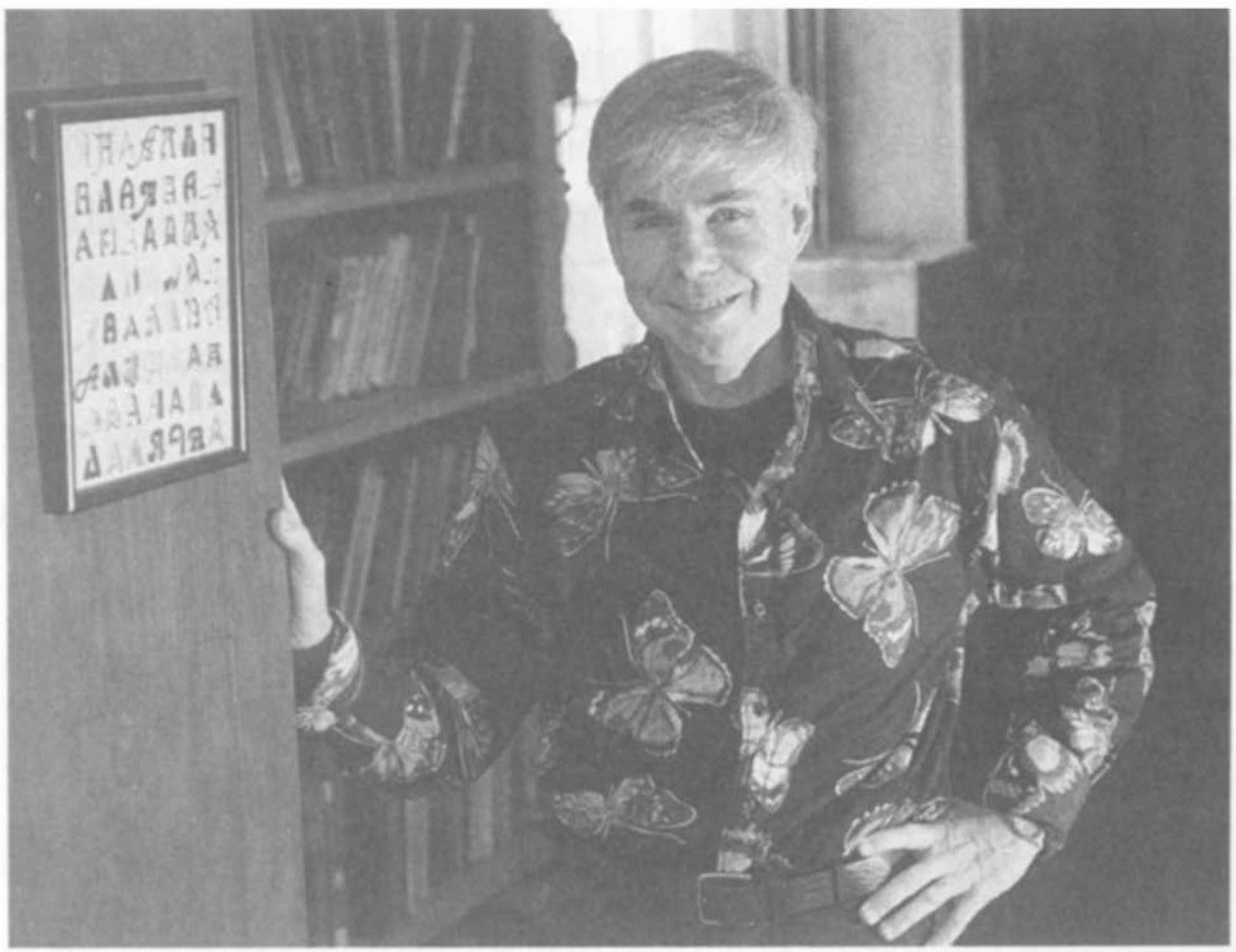


图13.1　侯世达（印第安纳大学提供照片）

问题是，我只是一个刚刚大学毕业的无名小卒，而侯世达则是获得了普利策奖（Pulitzer Prize）和美国国家图书奖（National Book Award）的著名畅销书作家。我给他写了一封信，说我想跟他读研究

生。但是我没有收到回信（后来才知道他没有收到那封信），因此我只好等待时机，并且学了一些AI的知识。

一年后我搬到了波士顿，换了工作，并学习了计算机课程，为我将来的事业做准备。有一天我碰巧看到侯世达将要到麻省理工学院（MIT）演讲的海报。这真让人兴奋，我立刻决定前往，挤进了狂热的书迷中（不仅仅只有我被侯世达的书改变了），希望能和他近距离接触。我终于挤到了前面，握到了侯世达的手，还告诉他我想参与他的AI研究，希望能申请印第安纳大学。他告诉我他实际上就住在波士顿，这一年他在MIT人工智能实验室访问。当时因为后面还有很多书迷在等着，侯世达就让我去旁边和他以前的一个学生详谈，转而接待其他读者。

我很失望，但是并没有放弃。我设法找到了侯世达在MIT人工智能实验室的电话号码，并且拨了几次。每次都是秘书接的电话，她告诉我侯世达不在，让我有事可以留口信。我留了口信，但是没有收到答复。

此后有一天晚上，我躺在床上琢磨该怎么办。我突然想到，我打电话的时候都是在白天，他都不在那里。既然侯世达白天总是不在，那他什么时候会在呢？肯定是在晚上！当时已是晚上11点，不过我还是起来拨了那个熟悉的号码。接电话的正是侯世达。

他很友好，和蔼可亲。我们谈了一会儿，他邀请我第二天去他的办公室谈，看我能在他的研究小组里做些什么。我如约而至，然后我们谈论了侯世达当时正在研究的课题——写一个能进行类比的计算机程序。

有时候要想有所收获，得有点斗牛犬的精神。

简化的类比

侯世达有一个天赋，他能将复杂的问题简化，然而又留住问题的精髓。在研究类比问题时，侯世达创造了一个微型世界，这个世界虽然是微型的，却保留了问题大部分有趣的方面。微观世界中包含在字母符号串之间进行的类比。

举个例子，思考下面的问题：如果abc变成abd，　ijk应该变成什么呢？大部分人会将变化描述为“将最右边的字母用其后继字母替换”，因

此答案是ijl。但其他答案也有可能，比如说：

◆ijd（“将最右边的字母用d替换”——就好像杰克将袜子“穿上”）

◆ijk（“将c用d替换；在ijk中没有c”）

◆abd（“不管什么字母串，都用abd替换”）

显然有无穷多种可能的答案，虽然可能性要小些，比如ijxx（“将c用d替换，将k用两个x替换”），但几乎所有人都认为ijl是最佳答案。不过这毕竟是个没有实际意义的抽象问题，因此如果你真觉得ijd好些，我也没法让你相信ijl更好。但是人类似乎进化出了在现实世界中进行类比的能力，以便更好地生存和繁衍，而他们的类比能力似乎也能应用于抽象领域。这意味着几乎所有人都会从内心同意有一个特定的抽象层次是“最合适的”，因而得出答案ijl。那些从内心会相信ijd是更好答案的人可能在进化过程中已经被淘汰了，这解释了为什么现在这样认为的人寥寥无几。

再来看第二个问题：如果abc变成abd，那iijjkk应变成什么？abckji⇒?abd仍然可以看作“将最右边的字母用其后继字母替换”，但如果将这条规则直接应用于iijjkk，得到的答案就是iijjkl，没有考虑到iijjkk的字母重复结构。大多数人会认同答案iijjll，背后的规则是“将最右边的字母组合用其后继字母的组合替换”，将abc中字母的概念变成了iijjkk中重复字母组合的概念。

在下面的问题中可以看到另一种概念迁移（conceptual slippage）：

abckji⇒?abd

kji⇒?

如果直接应用规则“将最右边的字母用其后继字母替换”，得到的答案就是kjj，但这样就没有考虑到kji的反向结构，kji是从右向左呈升序结构，而不是从左向右。这使得abc中的概念右迁移到了kji中的概念左，从而产生出新的规则，“将最左边的字母用其后继字母替换”，得出答案lji。大部分人都认同这个答案，有些人则倾向于答案kjh，将kji视为方向仍然是从左往右，只是采取的是降序。这里是将“后继字母”迁移为“前

继字母”，因此新规则就成了“将最右边的字母用其前继字母替换”。

再看下一个问题：

abckji⇒?abd

mrrjjjmrrjj⇒?

你想利用abc为字母升序这个明显的事实，但是现在呢？abc的内在结构很显眼，似乎是这个字母串的主要特征，但是mrrjjj似乎不容易看出有这样的结构。因此你可能会（像大多数人一样）认同mrrkkk（或是mrrjjk），也可能会多琢磨一下。这个问题的有趣之处在于，mrrjjj的背后正好潜藏着一个特征，认识到了这个特征，就能得出一个让大多数人更满意的答案。如果你忽略mrrjjj的字母，只注意其字母组合的长度，就能发现所期望的连续结构：字母组合的长度按“1—2—3”递增。一旦发现了abc和mrrjjj之间的这个关联，就可以得出规则“将最右边的字母组合在长度上增加一个”，在抽象层面上变成“1—2—4”，在具体层面上则对应为mrrjjjj。

最后再来看看下面的问题：

abckji⇒?abd

xyz⇒?

粗一看这个问题似乎与前面的ijk那个问题一样，可是有一个问题：z没有后继字母。大多数人的答案是xya，但是在侯世达的微型世界中，字母表不是循环的，因此这个答案不成立。这个问题陷入了僵局，进行类比的人需要重新审视他们最初的观点，可能需要原来没有考虑过的概念迁移，从而发现一种不同的方式来对问题进行理解。

人们对这个问题有各种答案，包括xy（“干脆将z去掉”），xyd（“将最右边的字母用d替换”；由于问题不寻常，这个答案虽然不那么严格，但是比前面的ijd要合理），xyy（“如果不能用z的后一个字母，那么就不如用它的前一个字母”），等等。然而有些人却似乎有天才般的洞察力，对这个问题能另辟蹊径。其中关键是注意到abc与xyz互为“镜像”——xyz位于字母表的末端，而abc则位于前端。因此xyz中的z

可以看作与abc中的a对应，很自然的x就与c对应。在这种对应背后是一组平行的概念迁移：字母表头kji⇒?字母表尾，最左kji⇒?最右，后继kji⇒?前继。这些迁移合在一起，就将最初的规则变成了适用于xyz的规则："将最左边的字母用其前一个字母替换"。从而得出很让人吃惊却又很有说服力的答案：wyz。

现在应该很清楚了，要在这个微型世界中进行类比，同在现实世界中一样，关键就是我所说的概念迁移。根据当前的背景找到合适的概念迁移对于找到好的类比极为重要。

模仿者

侯世达给我制订的计划是编写出能在字母串世界中进行类比的计算机程序，程序采用的机制要与人类进行类比的机制基本类似。他还为这个尚未存在的程序起了个名字："模仿者（Copycat）"。其中的思想是，类比是一种形式很微妙的模仿——例如，模仿abc转变成abd的方式，根据ijk自身的相关概念对其进行变化。因此这个程序是一位聪明而且富有创意的模仿者。

1984年夏天，我在MIT研究这个课题。到秋天，侯世达转到了位于安阿伯市（Ann　Arbor）的密歇根大学任教。我也跟着过去了，在那里攻读博士学位。我与侯世达一起花了近6年时间编写他构想的这个程序——当然，魔鬼藏在细节之中。终于，我们写出了能在微型世界中进行类似于人类的类比的程序，我也获得了博士学位。

如何做到

要想成为一个具有智能的模仿者，你首先必须理解你所"模仿的"对象、事件或情景。如果所面临的情景具有很多组成部分，各部分之间又有各种潜在的关系，比如一个视觉场景、一个朋友的故事或是一个科学问题，人（或者一个计算机程序）应该如何处理理解情景的大量可能方式以及与其他情景可能的相似性呢？

下面是两种相反的策略，都是不合理的：

1.有些可能性先天就被绝对排除。例如，在扫描mrrjjj之后，列出一

系列可能的概念（比如字母、字母组合、后继、前继、最右等等），然后严格限定于此。显然，这个策略的缺点是放弃了灵活性。一些与问题的关联不那么明显的概念可能后来发现是关键性的。

2.所有可能性都是平等的，处理的难度也是一样的，因此可以穷尽搜索所有可能的概念和关系。这种策略的问题是在现实世界中有着太多的可能性，甚至事先不知道对于给定情形有哪些可能的概念。如果你的发动机发出不太正常的突突声，你的汽车无法发动，你可能会认为有两种原因，可能性一样，（a）皮带从轴承上脱落了，或者（b）皮带老化，断了。如果没什么特别的原因你就认为还有一种同等的可能性是你的邻居偷偷把你的皮带剪了，那你是有点妄想迫害症。如果你还认为有种同等的可能性是组成皮带的原子通过量子隧道进入了另一个平行宇宙，那你是科学狂人。如果你不断想出各种各样同等的可能性，那……正常人的头脑是没法做到的。不过，也许你刚好就有一个恶毒的邻居，量子隧道的可能性也不能完全排除，说不定你还会因此获得诺贝尔奖，这些可能性都是存在的。

所有可能性都有可能存在，但问题的关键是它们的可能性不是对等的。不符合直觉的可能性（例如，你恶毒的邻居、量子隧道等等）确实有可能，但必须有明确的原因才会被加以考虑（例如，你听说过你邻居的恶行；你刚好在你的车上安装了一台量子隧道机器；其他可能性被发现是错误的）。

大自然在许多例子中都找到了探索性策略来解决这个问题。例如，在第12章我们看到蚁群搜索食物的方式：通往最佳食物源的最短路径会具有最强烈的信息素气味，沿路径行进的蚂蚁越来越多。然而，无论何时，还是会有一些蚂蚁沿着气味较弱、可能性不大的路径前进，也有一些蚂蚁仍然会随意搜寻，这样就有可能发现新的食物源。

这是一个在搜索和开发之间进行平衡的例子，在第12章曾提到过这一点。如果收益很可观，就必须根据估计的收益以一定的速度和强度进行开发，并不断根据新的情况加以调整。但无论何时都不停止对新的可能性的探索。问题是如何根据最新的信息为各种可能动态分配有限的资源——蚂蚁、淋巴细胞、酶或者思维。蚁群的解决方案是让大部分蚂蚁采取两种策略的组合：不断随机搜索与简单地跟随信息素轨迹并沿途留下更多信息素的反馈机制相结合。

免疫系统在探索和开发之间似乎也保持着几乎最优的平衡。在第12章我们看到免疫系统利用随机性获得了对遇到的任何病原体做出反应的潜力。一旦抗原激活了某个B细胞，触发了这种细胞的增殖，潜力就会被释放出来，繁殖出的抗体会增加针对这种抗原的特异性。因此免疫系统对所发现的抗原信息进行开发的方式就是分配大量资源来对抗目前发现的抗原。但是它仍然会保有大量不同的B细胞，以继续探索其他可能性。同蚁群一样，免疫系统也是将随机性同基于反馈的高度导向行为结合在一起。

侯世达提出了一种探索不确定环境的方案，“并行级差扫描”，在第12章我曾提到过。根据这个方案，许多可能性被并行地进行探索，用获得的最新信息不断对各种可能性的收益进行估计，并根据反馈分配资源。同蚁群和免疫系统一样，所有可能性都有可能被探索，但是在同一时刻只有部分被探索，并且分配的资源也不一样多。当人（或蚁群，或免疫系统）对所面临的情形只有很少的信息时，对各种可能性的探索开始时非常随机、高度并行（同时考虑许多可能性）和分散：没有理由要特别考虑某种可能性。随着获得的信息越来越多，探索逐渐变得集中（增加的资源集中于少数可能性）和确定：确实有收益的可能性会被开发。同蚁群和免疫系统一样，在模仿者（copycat）中，这种探索策略也是通过简单个体的大量互动涌现出来。

模仿者程序

模仿者的任务是用其拥有的概念在某个问题的三个未经处理的字符串之上建构认知结构：对对象的描述、相同字符串中对象的关联、字符串中对象的分组以及不同字符串中对象的对应关系。程序建立的结构代表其对问题的理解，让它可以得出一个答案。对于每个问题，程序的初始状态都是一样的，拥有的概念集也完全一样，因此概念必须具有适应性，根据其相互之间的关联以及不同问题的情形进行适应。面对问题时，建立起对情形的表示之后，联想就会产生出来，并且被视为各种可能性，并行级差扫描以并行的方式对理解情形的各种可能途径进行检验，并根据当前对各种可能性的收益评估决定对其检验的速度和深度。

模仿者对字符串类比问题的解答涉及以下这些模块的互动：

◆移位网（Slipnet）：概念组成的网络，包含一个中心节点，周围

围绕着可能的联想和移位。图13.2中给出了程序最新版中一些概念和关系的图。移位网中每个节点都有一个动态的活性值，表示目前认识到的其与正在处理的类比问题的关联性，这个值随着程序的运行不断调整。活性值会向相邻概念扩散，并且如果得不到加强就会衰减。每条连线有一个动态的阻抗值，表示其目前对移位的阻力。这个值也随着程序运行不断变化。连线的阻抗值与对应节点的活性值呈反比。例如，当反向活性很高时，由反向连线相连的节点（如后继和前继）的移位阻抗值就很低，从而增加移位的可能性。

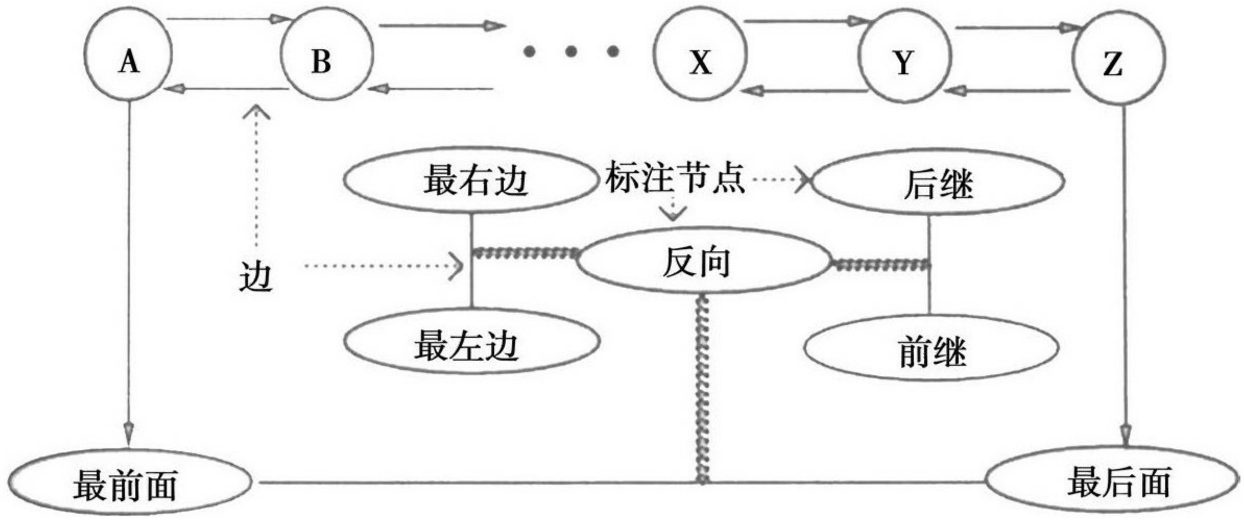


图13.2　模仿者的移位网局部。节点用其表示的概念标注（例如，A—Z、最右边、后继）。一些节点之间的连线（例如最右边—最左边）连接到一个节点，表示连线所代表的关系（例如反向）。每个节点都有动态的活性值（没有标出来），活性会向相邻节点扩散。如果得不到增强，活性值就会衰减。每条连线都有移位阻抗，对应节点活性越高，连线阻抗就越低

◆工作区：作业的区域，其中有类比问题的字母和在字母上建立的认知结构。

◆码片（Codelets）：在工作区不断探索可能认知结构的自主个体，自主个体会试图实现它们发现的结构。（码片一词意指“小片的编码”。）成组的码片一起合作构建出定义对象关系的认知结构（例如，“abc中a的后面是b”，或“iijjkk中的两个i形成组合”，或“abc中的b对应于iijjkk中的jj”，或“abc中的c对应于kji中的k”）。每组码片考虑世界结构的一种特定可能性，根据试图建立的结构的可能收益为各组分配资源（码片时间），可能的收益随着探索的进行不断进行评估。这样就实现了对各种可能性的并行级差扫描，各组码片通过竞争和合作，逐步建立起结构层次，表现程序对情形的“理解”。

◆温度：对系统的认知组织程度的度量。类似于物理世界，温度高对应无组织，温度低对应高的组织度。在模仿者中，温度度量组织程度，并作为反馈信号控制码片决策时的随机程度。温度高时，表明认知组织度低，能据以进行决策的信息也少，码片决策时也就越随机。随着认知结构的建立，对于相关的概念以及如何建构对世界中对象和关系的认知的信息也越来越多，温度也就越来越低，表明有更多的信息引导决策，码片进行决策时也就越有确定性。

运行模仿者

解释模仿者各部分之间如何交互的最好方式是用图形显示程序的运行情况。这些图形是程序运行时的屏幕显示。这一节我们来看一下程序面对abckji⇒?abd, mrrjjjmrrjj⇒?这个问题时的运行过程。

图13.3：给出的问题。图中包括：工作区（这里是类比问题中尚未结构化的字母）；左边竖条为“温度计”，显示当前的温度（初始值设为100，100也是最大值，反映出当时没有任何认知结构）；右下角显示当前运行的码片数量（初始值为0）。

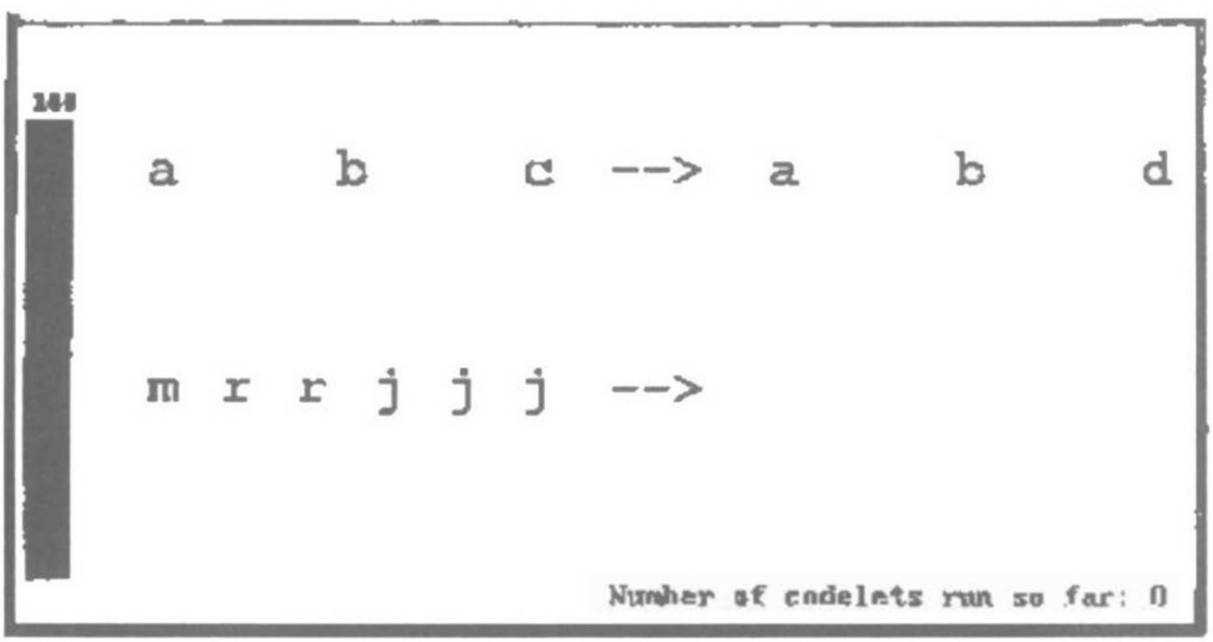


图13.3

图13.4：有30个码片在运行，已经探索了许多可能的结构。原则

上，码片可以视为同蚂蚁一样的自主体，每个都根据一定的概率探索一条路径，但是受其他码片的探索路径引导。在这里“路径”代表的是可能的认知结构。码片随机搜索合理的描述关系、字符串划分以及字符串之间的对应关系，然后提出可能的结构。如果认可某种结构的码片很多，这种结构就会得到增强。一旦强度达到某个阈值，就认为结构被建立起来了，从而影响后面结构的建立。

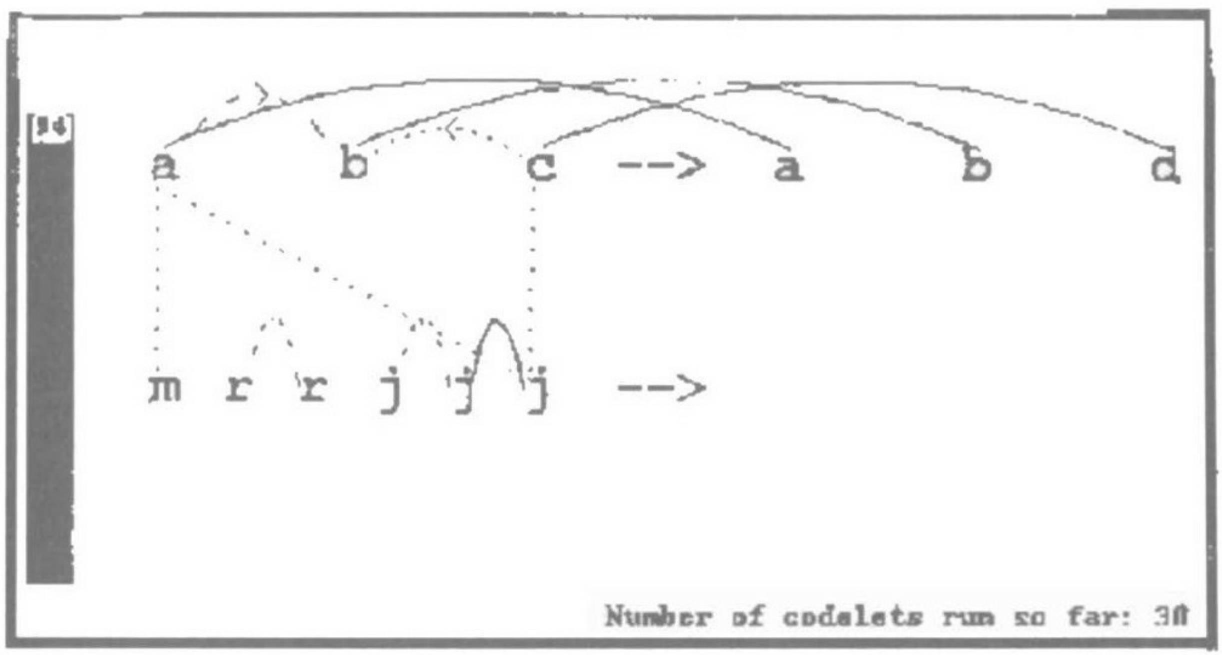


图13.4

在图13.4中，点线代表刚刚开始考虑的结构；短画线代表考虑了一段时间的结构；实线代表已经建立起来的结构。码片根据对结构可能收益的判断决定提出这种结构的速度，例如，提出a—m对应关系的码片认为其有很高的可能收益，因为两者都在各自字符串的最左边：最左边kji⇒?最左边这样的一致性关系总是很强。提出a—j对应关系的码片则认为其要弱得多，因为最左边kji⇒?最右边这样的对应要弱很多，而且反向关系目前也没有活性。因此对a—m对应关系的探索速度就很可能比不太合理的a—j对应要快得多。

因为mrrjjj中最右边的两个j建立起了“相同”关系，温度值从100降到了94。这个相同关系激活了移位网中的节点相同（没有画出来）。这会驱使一些码片去寻找其他相同的地方。

图13.5：有96个码片在运行。abc的后继关系已经建立起来。13.4中提出的c—b的前继关系已经替换为b—c的后继关系。abc中的两个后继关系相互支持：每个都因为对方的存在而变得更强，因而使得与之竞争的前继关系不太可能胜出。

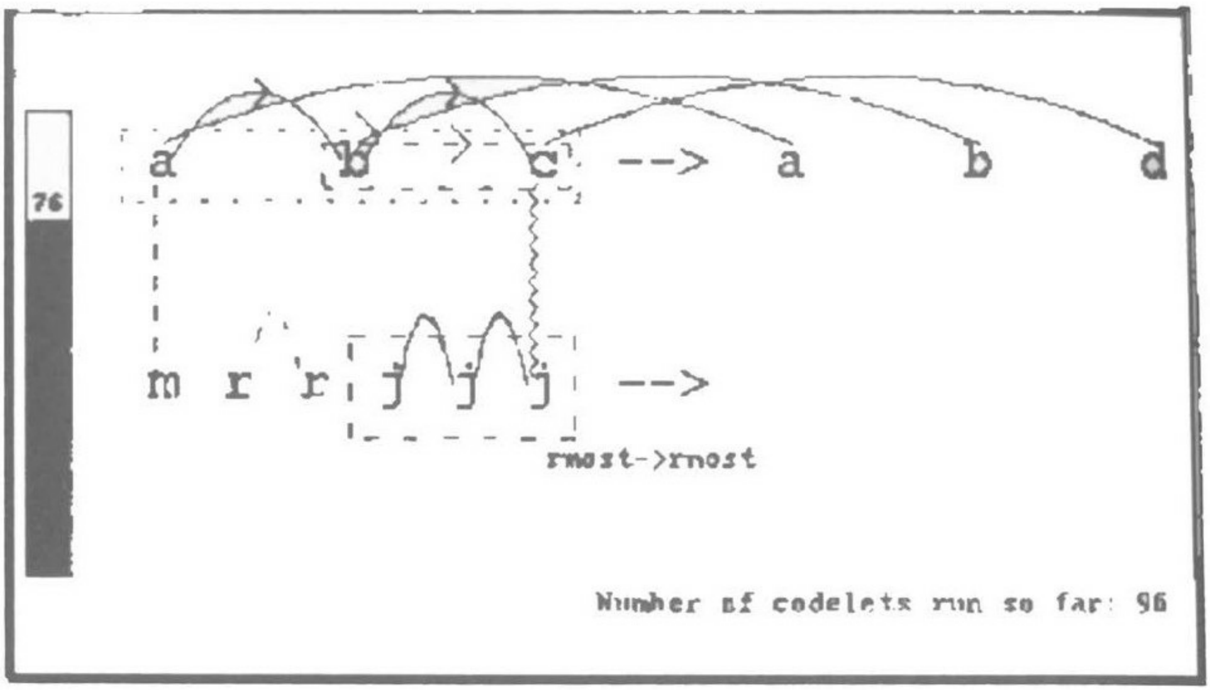


图13.5

两个都是基于后继关系的组合相互竞争：bc和abc（整个字符串形成组合）。两者在图13.5中分别用点线方框和短画线方框表示。虽然bc组合出现的时间早些（它是短画线，而后者是点线），abc组合涵盖的字母却更多。因此它比bc组合要强些——码片会倾向于以较快的速度对其进行验证，也更有可能将其的结构建立起来。图下方在考虑基于相同关系的组合jjj。

a—j的交叉对应关系（图13.4中的点线）已经被放弃了，因为进一步探索它的码片发现其太弱，无法建立起来。c—j对应则被建立起来了（垂直的锯齿线）；线的下端给出了对应关系的依据（两者都是各自字符串最右边的字母）。

因为后继和相同关系都建立起来了，再加上最右边⇒最右边（rmost⇒rmost）对应，这些节点在移位网中活性很高，因而会驱使码

片去探索其他地方是否还有这种关系。例如，考虑最左边字母的对应关系。

随着结构的建立，温度也相应降到了76。温度越低，码片的决策就越不随机，因此像bc组合这样不太可能的结构就更不可能被建立起来了。

图13.6：abc组合和jjj组合建立起来了，用字母周围的实线方框表示。为了让图像清晰，组合中字母之间的连接没有显示。这些组合的存在驱使码片寻找其他后继和相同关系的组合，例如rr相同组合就被高度关注。jjj这样的组合形成了字符串中新的对象，可以有它们自己的描述，有自己的连接，也可与其他对象产生对应。大写字母J表示由jjj组合形成的对象；类似的，abc组合也形成一个新的对象，为了清晰起见，图中没有画出表示它的单个字母。同其他特性一样，组合的长度并不能被程序自动注意到，要由码片来发现。每当一个组合节点（例如，后继组合、相同组合）在移位网中被激活，它就会将部分活性扩散到长度节点。这样长度节点就有了一点活性，从而产生出关注长度的码片，但这些码片并不会马上就会相互比较，甚至还没有开始运行和关注组合的长度。

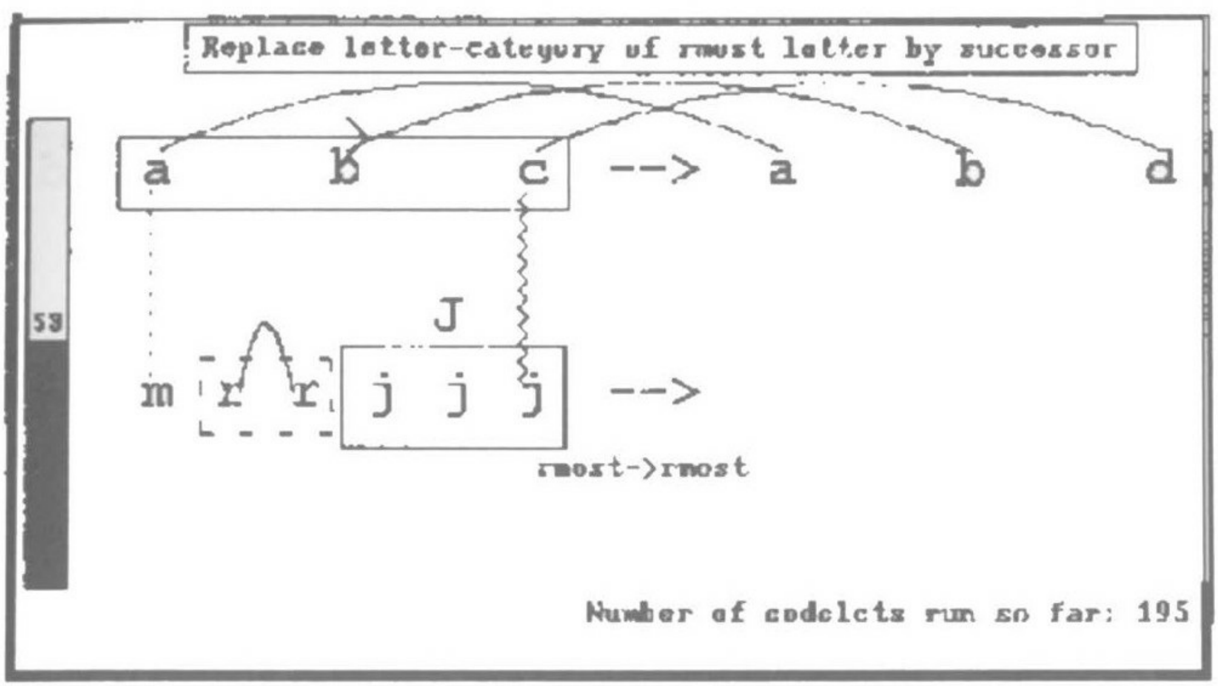


图13.6

一个描述abc⇒abd变换的规则已经建立起来了："将最右边的字母单元用其后继字母替换"。模仿者当前的版本认为变换示例替换的刚好是一个字母，因此建立规则的码片会对模板"将___用___替换"进行填充，以一定的概率在程序给出的字母变换描述中进行选择，描述越抽象（例如最右边的字母就比C抽象），选中的概率就越大。

温度降到了53，这是因为建立起来的结构表明认知组织度越来越高。

图13.7：有225个码片在运行。字母与字母的对应c—j已经被字母与组合的对应c—J代替。在对应关系的下方，最右边⇒最右边（rmost⇒rmost）上面增加了字母⇒组合（let⇒group）对应关系。c—J对应要强于c—j对应，因为前者涵盖的对象更多，同时组合概念的活性很高，因此与问题的相关度似乎更高。然而，虽然c—j对应相对较弱，还是有一组新的码片又在考虑它。

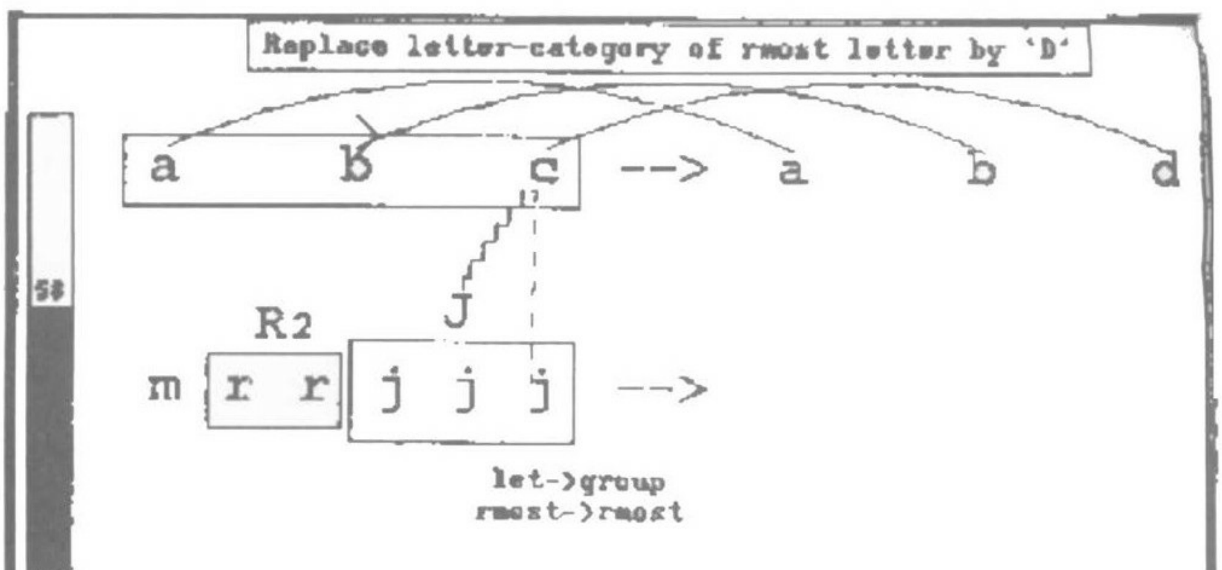


图13.7

与此同时，rr组合也建立起来了。此外，它的长度（用R后面跟着的2表示）也被一个码片注意到了（概率事件）。这个事件激活了长度节点，驱动码片注意其他组合的长度。

屏幕上方产生了新的规则，“将最右边的字母单元用‘D’替换”，虽然这条规则比前面被换掉的规则弱些，但由于各种结构（包括规则）的竞争是根据概率随机决定的，因此这条规则还是有机会胜出。然而它的胜出导致温度又回升到了58。

如果程序现在停止（这不太可能发生，因为决定程序是否停止的概率中有一个很重要的因素就是温度，而目前温度还很高），对于字符串mrrjjj，得出的规则就是“将最右边的字母组合用‘D’替换”因为要遵守根据c—J对应得出的字母⇒组合移位，从而得到答案mrrddd，模仿者程序确实会得出这个答案，不过次数很少。

在整个运行期间不断会有码片想要制造出答案，不过不太可能成功，除非温度很低。

图13.8：有480个码片在运行，“将最右边的字母单元用其后继字母替换”这条规则又恢复了。但是c—J对应没了，换成了c—j对应（也是一个概率事件）。如果程序在现在停止，得到的答案将会是mrrjjk，因为abc中的c对应的是一个字母，而不是组合。这样建立答案的码片就会忽略b与字母组合的对应。不过c与组合J的对应再次受到强烈关注。它将与c—j对应竞争，不过它将比以前更强大，因为与之平行的b与组合R的对应已经建立起来了。

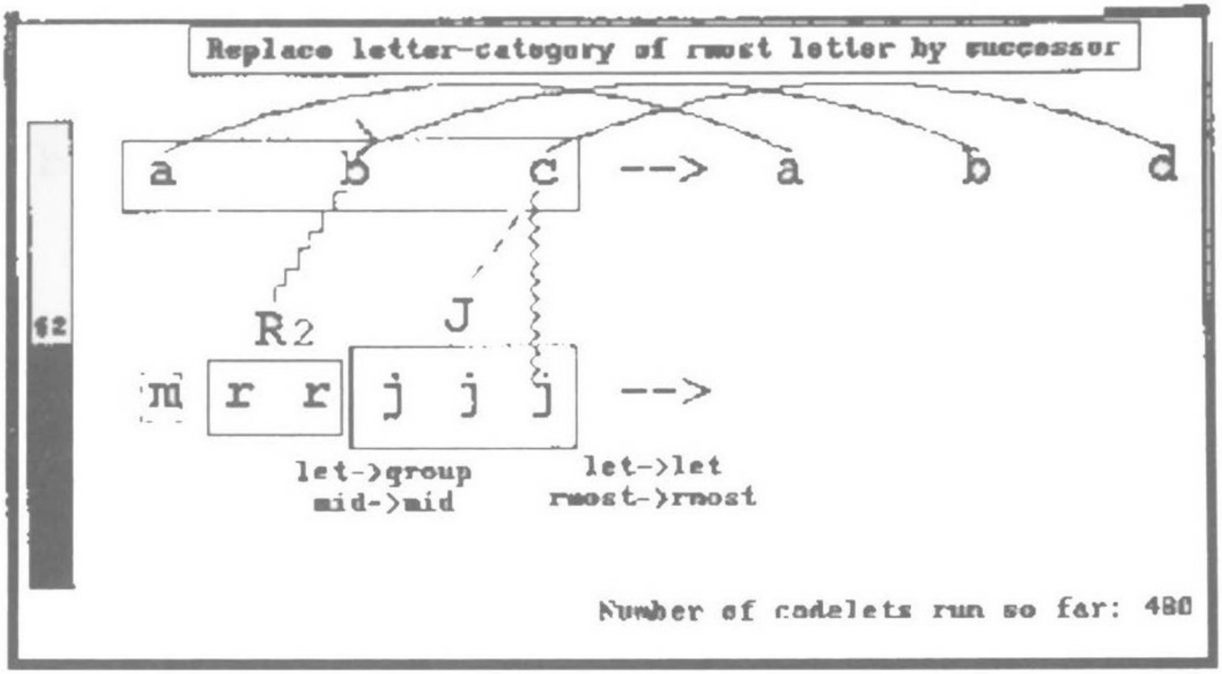


图13.8

在移位网中，长度的活性衰减了，因为对R组合的长度描述目前还没有发现有什么用处（也就是说，没有发现其与其他结构的关联）。在工作区中，描述R的长度的2已经消失了。

温度仍然很高，因为目前程序还很难为mrrjjj建立一个一致的结构，为abc建立结构则容易些。由于问题一直没有解决，再加上在mrrjjj中建立的两个相同组合导致的驱动力，使得系统开始考虑单个字母组成的相同组合这种不太可能出现的结构。这表现为围绕着字母m的短画线方框。

图13.9：这些驱动力结合到一起，使得相同组合M被建立起来，与R组合和J组合并列。其长度1附在了其描述后面，激活了长度节点，使得程序再次考虑组合长度与问题有关联的可能性。这次激活强烈地吸引了码片关注组合的长度。一些码片已经在探索它们之间的关联，而且在1和2之间建立了后继关系，这可能是受abc中的后继关系的驱使。

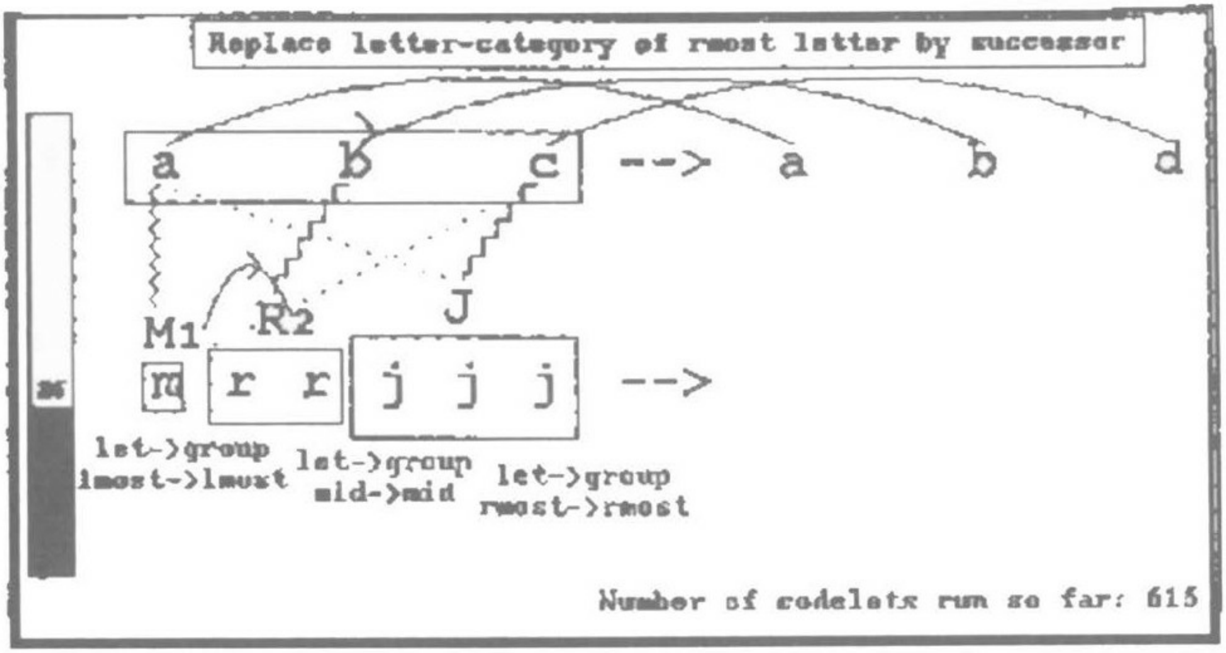


图13.9

三个一致的字母⇒组合（let⇒group）对应已经建立起来了，由于这个新结构的前景被看好，温度也降到了36，反过来又使得这个新观点被锁定。如果程序这个时候停止，它将得到答案mrrkkk，这也是最常见的答案（图13.13）。

图13.10：由于长度节点一直具有活性，其余两个组合（jjj和abc）也附上了长度描述，2和3之间的后继关联也被注意到了（很大程度是受abc和mrrjjj的结构驱动）。不太可能的候选结构（bc组合和c—j对应）仍然被留意，不过不像以前那样关注了，现在对问题的一致认识开始涌现出来，温度也相当低了。

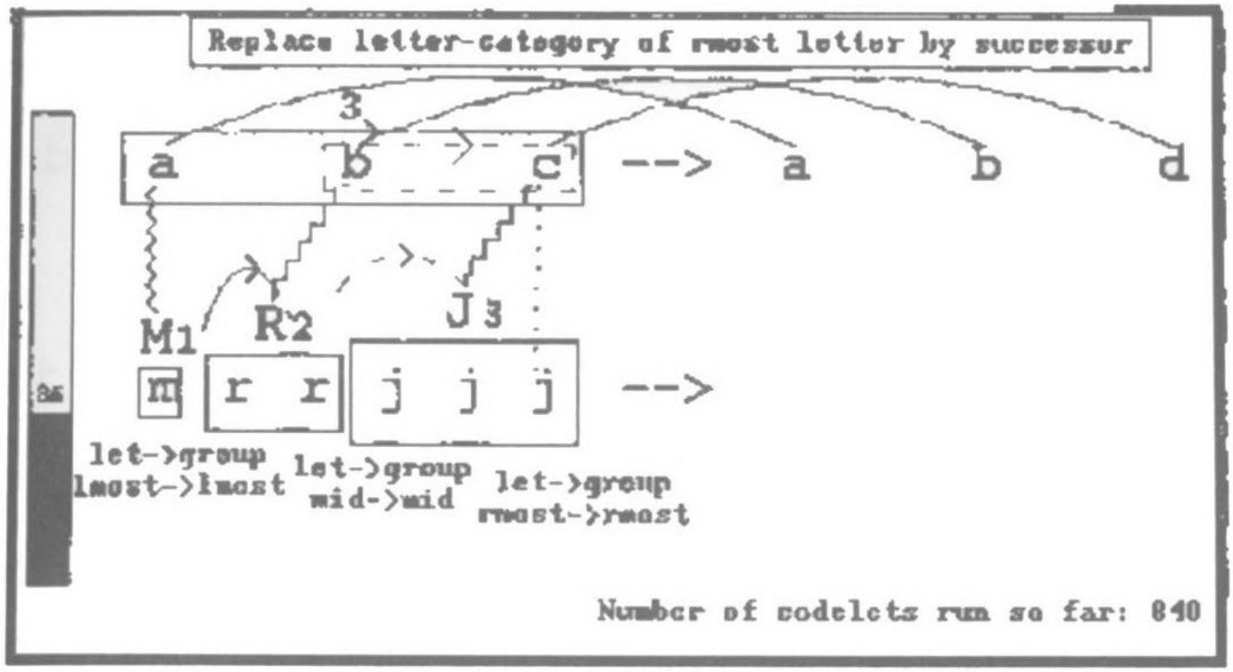


图13.10

图13.11：2和3之间的关联建立起来了，再加上abc后继组合的驱使，使得码片能够基于数字（而不是字母）后继关系提出并建立由整个字符串构成的组合。这个组合用围绕三个相同组合的实线框表示。同样，由abc和mrrjjj整个字符串构成的这两个组合之间的对应（两个字符串右边的垂直点线）也受到关注。

可笑的是，这些复杂的思想似乎正要成形，少数逆潮流而动的码片叛徒却撞了好运：它们成功地扳倒了c—J对应，代之以c—j对应。显然这降低了整体水平；强大的mrrjjj组合的建立本来能让温度降低许多，然而这种降低却被两个字符串之间对应关系的不平行抵消了。如果这时强行停止程序运行，得出的答案将是mrrjjk，因为这时c似乎更倾向于与字母j而不是组合J对应。不过另外两个对应关系会继续驱使程序（作用在码片上）回到c—J对应。

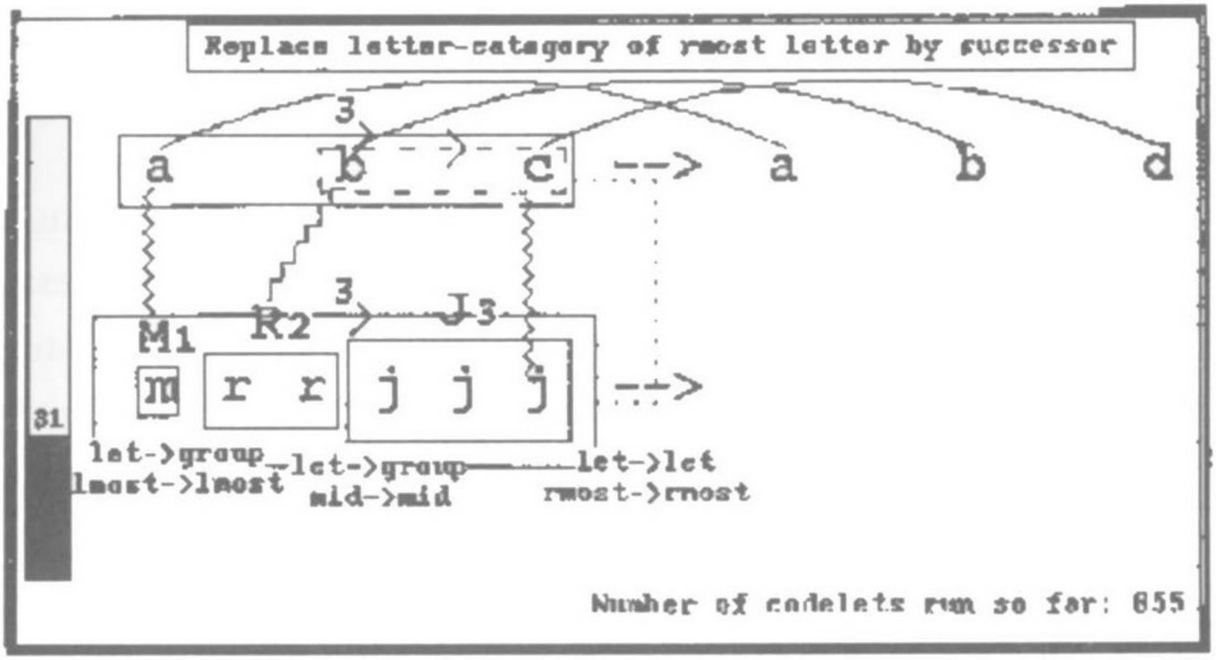


图13.11

图13.12：事实上，没运行多久就发生了这种情况：c—j对应被打破，c—J对应重新建立起来。另外，abc与mrrjjj整个字符串之间的对应也建立起来了；上面标着整体⇒整体（whole⇒whole），后继组合⇒后继组合（sgroup⇒sgroup），右⇒右（right⇒right，组合中连接的方向），后继⇒后继（successor⇒successor，组合中连接的类型），字母⇒长度（letcat⇒length），以及3⇒3（组合的长度）。

现在程序建立的认知结构已经非常一致了，使得温度降得很低（11），由于温度低，一个码片已经成功根据工作区中给出的移位（字母⇒组合，字母⇒长度，其他对应都是不变对应）得出了规则。得出的规则就是“将最右边的组合的长度用其后继进行替换”，从而得出答案mrrjjjj。

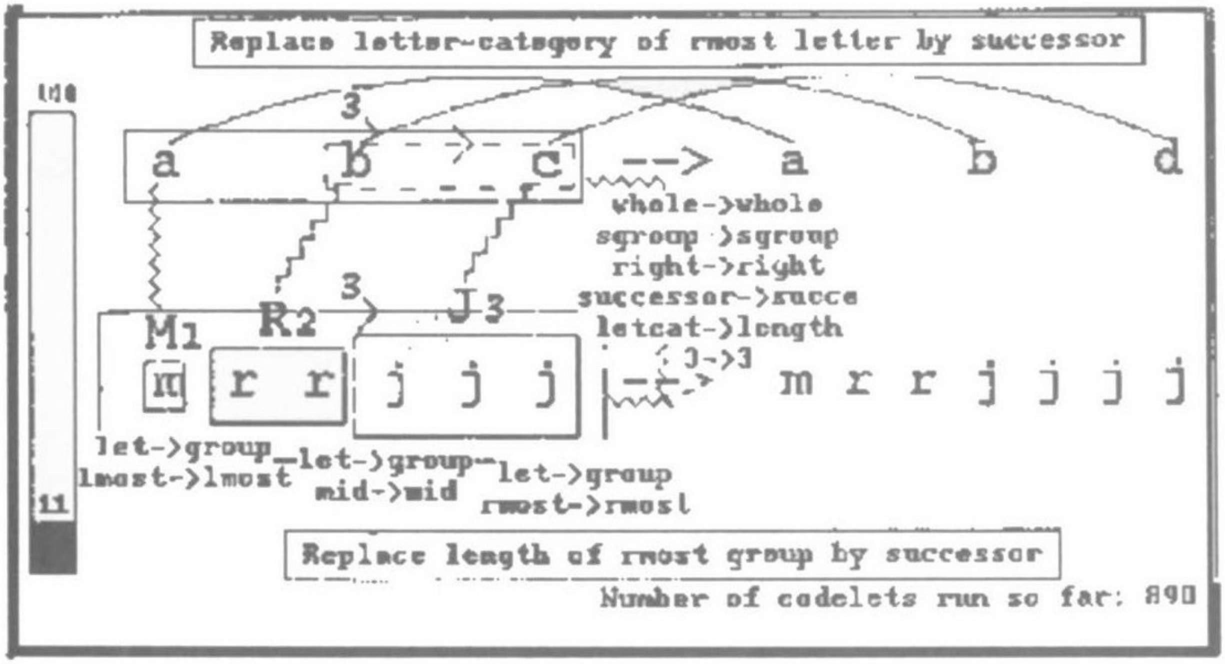


图13.12

从上面的叙述可以清楚地看到，由于模仿者每次运行都充满了随机决定，因此运行得出的结果也可能会不一样。图13.13显示了模仿者1000次运行中给出的不同答案，每次运行都有不同的随机数种子。各种答案的立方条高度表示答案的出现频率，上面标了具体的次数。立方条下面则给出了各种答案最后的平均温度。

答案的频率大致上对应于答案相对于程序的显见程度。例如，对于程序来说，出现了705次的mrrkkk就比只出现了42次的mrrjjjj要显见得多。但是得出mrrjjjj后的平均温度还是比得出mrrkkk后的平均温度要低得多（分别是21和43），这表明虽然后者更显见，程序还是认为，根据答案结构的一致性，前者要好于后者。

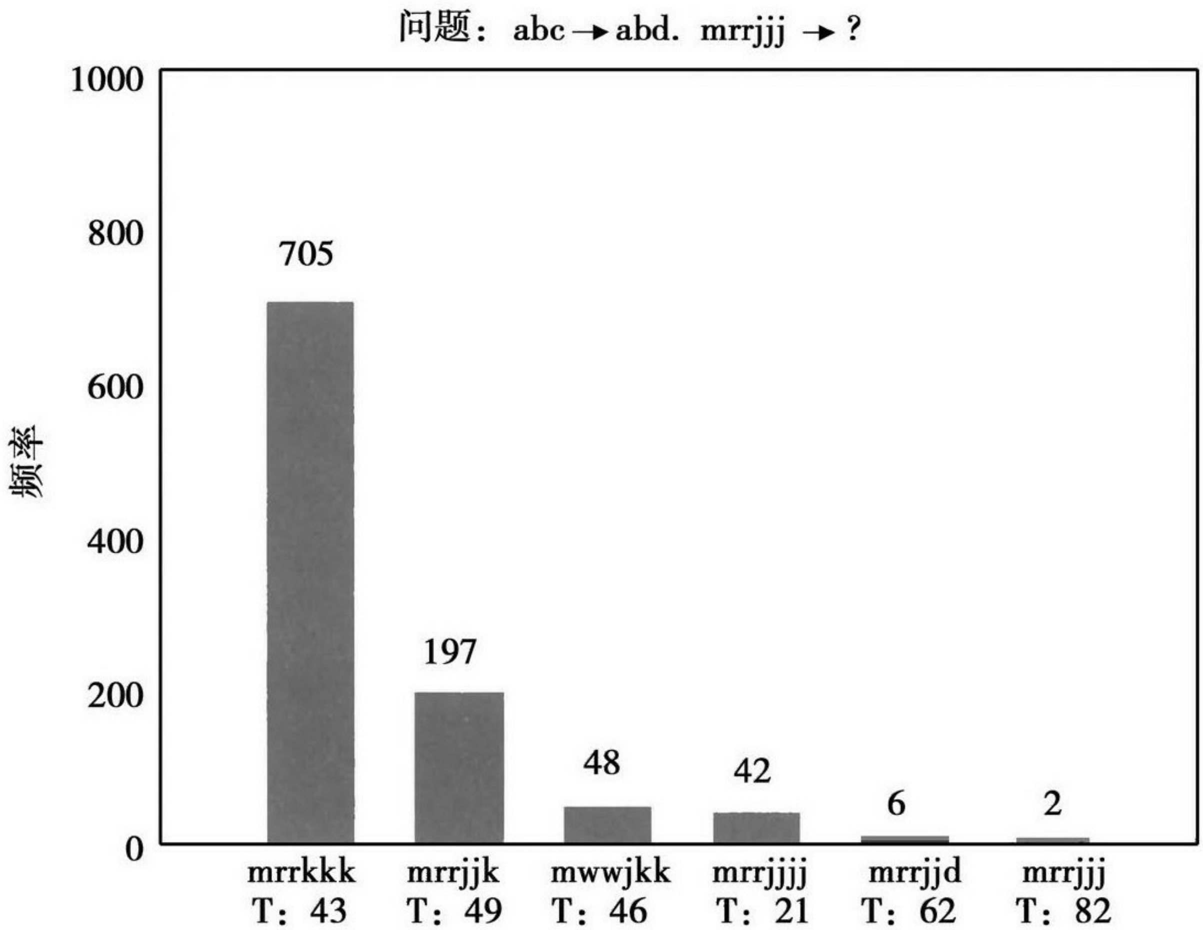


图13.13让模仿者运行1000次，每次的随机数种子都不相同，得出的最后答案的样条图

总结

通过程序运行中表现出来的机制，模仿者避开了认知的第22条军规：你无法探索所有可能，但如果你不探索它们，你就无法知道哪种可能值得探索。你必须毫无偏见，但需要探索的领域又太大；你需要利用概率来让探索公平。在模仿者受生物启发的策略中，最初是信息很少，使得温度很高，也很随机，许多探索并行进行。随着获得的信息越来越多，适当的概念被发现，温度降低，探索也变得越来越具有确定性，一系列特定概念开始占据主导。整体上使得系统从极为随机、并行和自底向上的处理模式逐渐转变成确定、连贯而且集中的模式，逐步发现对情形的一致性认知并渐渐让其“凝结”。就像我在第12章阐释的，处理模式的逐步转变似乎是一些复杂适应系统的共性。

模仿者与生物系统的相似性迫使我们从更广泛的角度来思考和理解

我们所建立的系统。例如，人们可能会注意到，细胞因子在免疫系统中传递信号的作用就类似于码片在类比问题中唤起对特定位置的注意，这实际上就是在广义的信息处理层面上认识生物学单位的功能。类似的，如果注意到了免疫系统中的温度现象——发热、炎症——来自于许多自主体的联合行动，人们也许就会受到启发，从而更好地对模仿者这样的系统中的温度机制进行建模。

最后还有棘手的意义问题。在第12章我曾说过，就传统计算机来说，信息对于计算机本身是没有意义的，只对人类创造者和“最终使用者”才有意义。然而，我还是认为模仿者代表了一种相当不传统的计算模型，对于它所具有的概念和它所做的类比，它能理解一种非常初级的意义。例如，后继组合的概念在网络中与其他类似的概念相联，模仿者能够认识这些概念，并在各种情形中合理地利用这些概念。在我看来，这就是意义的开端。但是就像我在第12章说的，意义与生存和自然选择密不可分，而模仿者还没有涉及这一些，只是在降低温度这一点上与“生存”本能有一点点关联。从这方面来说，模仿者（以及后来由侯世达的研究团队研发的一系列让人印象更加深刻的程序[187]）与生物系统还有着很大的区别。

AI的终极目标是让人摆脱意义的怪圈，并且让计算机本身能理解意义。这是AI中最难的问题。数学家罗塔（Gian-Carlo Rota）称这个问题为“意义屏障”，[188]不知道AI是否或何时能“破解”它。我个人认为这不是短时间能做到的，但如果能够破解，我怀疑类比将是关键。

第14章　计算机模型

复杂系统顾名思义就是很复杂的系统，而物理、化学、数学、生物学这些以数学为导向的学科关注的往往是易于用数学处理的简单而理想化的系统。复杂系统很难单独用数学进行处理，不过现在的计算机速度越来越快，价格也越来越便宜，已经有可能构造复杂系统的计算机模型并进行实验。图灵、冯·诺依曼、维纳（Norbert Wiener）等计算机科学先驱都希望用计算机模拟能发育、思维、学习和进化的系统。一门新的实践科学由此诞生。在理论科学和实验科学之外又产生了一个新的门类：计算机仿真（图14.1）。在这一章，我们来看一看复杂系统的计算机模型能告诉我们什么，用这样的模型来进行研究又会有哪些陷阱。

模型是什么

在科学中，模型是对某种“实在”现象的简化表示。科学家们说是在研究自然，但实际上他们做的大部分事情都是在对自然进行建模，并对所建立的模型进行研究。

以牛顿的引力定律为例：两个物体之间的引力正比于它们质量的乘积。这是对一种特定现象的数学描述——也就是数学模型。还有一种模型是用较为简单的概念来描述现象实际是如何运作的，也就是所谓的原理。在牛顿的时代，他的引力定律受到质疑，就是因为他没有解释引力的原理。也就是说，他没有用“大小、形状和运动”等物理对象的属性对其进行解释——根据笛卡儿的思想，这些基本要素是所有物理模型必要而且充分的组成部分[189]牛顿自己推测过引力的可能原理，例如，他“猜想地球就像海绵一样，[190]不断吸收天空降落下来的轻质流体，这种流体作用到地球上的物体上，导致它们下降”。这种概念框架可以称为原理模型。200年后，爱因斯坦提出了一种不同的引力原理模型——广义相对论，在其中引力被概念化为四维时空的几何特性。现在，一些物理学家又在鼓吹弦论，提出引力是由细小、振动的弦导致的。

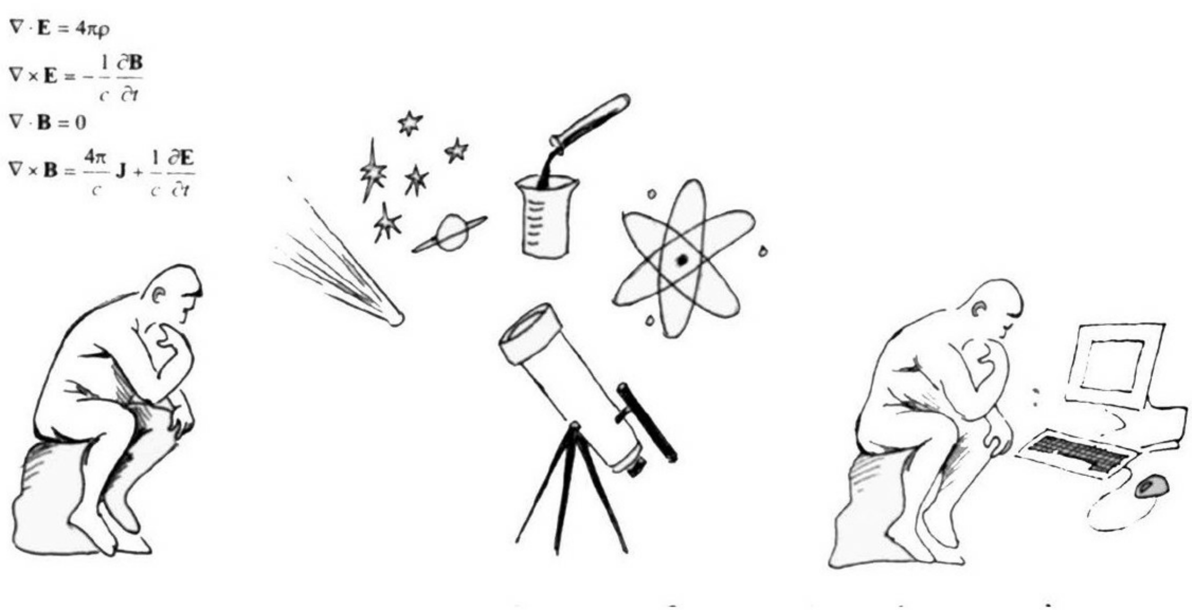


图14.1　在理论科学和实验科学的传统划分之外又产生了一个新的门类：计算机仿真（David Moser绘制）

模型是我们思维的方式，是用我们熟悉的概念解释观察到的现象，

所用到的概念是我们的头脑能够理解的（就弦论来说，则是少数非常聪明的人能够理解的）。模型也是预测未来的途径：比如说，牛顿的引力定律仍然被用来预测行星轨道，而爱因斯坦的广义相对论则成功预测了那些所预测的轨道的偏差。

理想模型

在天气预报、汽车和飞机设计、军事运筹中，经常用计算机来运行详尽而复杂的模型，对所建模的特定现象进行详细的预测。

而在复杂系统研究中一个主要的方向就是研究理想模型：通过相对简单的模型来理解一般性的概念，而不用对具体系统进行详细的预测。下面是我在书中曾讨论过的一些理想模型的例子：

◆麦克斯韦妖：用来研究熵的概念的理想模型。

◆图灵机：用来对“明确程序”进行形式化定义以及研究计算概念的理想模型。

◆逻辑斯蒂模型和逻辑斯蒂映射：用来预测种群数量的极简模型；后来成为研究动力学和混沌一般性概念的理想模型。

◆冯·诺依曼自复制自动机：用来研究自复制“逻辑”的理想模型。

◆遗传算法：用来研究适应性概念的理想模型。有时候也作为达尔文进化的极简模型。

◆元胞自动机：用于研究一般性的复杂系统的理想模型。

◆科赫曲线：用来研究海岸线、雪花等分形结构的理想模型。

◆模仿者：用来研究人类类比思维的理想模型。

理想模型有许多用途：研究一些复杂现象背后的一般机制（例如，冯·诺依曼研究自复制的逻辑）；证明解释某种现象的机制是不是合理（例如，种群数量的动力学）；研究简单模型在变化后的效应（例如，研究遗传算法的变异率或逻辑斯蒂映射的控制参数R变化所带来的影

响）；或者更普遍是作为哲学家丹尼特（Daniel Dennett）所谓的“直觉泵”[191]（intuition pump）——用来引导对复杂现象进行理解的思维实验或计算机仿真。

复杂系统的理想模型也能为新的技术和计算方法提供灵感。例如，图灵机启发了可编程计算机；冯·诺依曼的自复制自动机启发了元胞自动机；达尔文进化、免疫系统和昆虫社会的极简模型分别启发了遗传算法、计算机免疫系统和“群体智能（swarm intelligence）”方法。为了领略科学中理想模型的成就，现在我们来深入了解一下社会科学中的几个理想模型，从最广为人知的模型开始：囚徒困境。

对合作的进化进行模拟

许多生物学家和社会学家都用理想模型来研究为什么在由自私个体组成的群体中会进化出合作。

生物归根结底都是自私的——它们要想在进化中获得成功，就必须能活足够长的时间，保持足够的健康，还要能吸引异性，以繁衍后代。大部分生物为了达到这些目的会毫不犹豫地与其他生物进行斗争，采用各种伎俩，杀死或杀伤其他生物。通常的看法认为进化选择会使得自私或自卫本能得以传递给下一代并在种群中扩散。

然而与这种看法相反，在生物王国和社会的各个层面上都有许多明显不符合自私原则的例子。从底层看，在进化历程的一定阶段时刻，单细胞生物会互相合作以形成更复杂的多细胞生物。后来，又进化出了蚁群这样的社会性生物，大部分蚂蚁为了蚁群的整体利益工作，甚至放弃了繁衍的能力，只让蚁后来繁衍后代。后来，灵长类动物群体中又涌现出了更加复杂的社会，社会团结一致对外，复杂的贸易，最终出现了人类国家、政府、法律和国际条约。

生物学家、社会学家、经济学家和政治学家都面临着类似的问题，本质上自私的个体中是怎么产生出合作的。这不仅仅是个科学问题，也是政治问题：例如，是否有可能创造条件让国家之间产生并维持合作，一起应对核扩散、艾滋病、全球变暖等国际问题？

囚徒困境

20世纪50年代冷战高峰时期，一些人开始思考如何推动敌对国家之间的合作，以避免核战争。1950年前后，两位数学博弈论学家，弗勒德（Merrill Flood）和德雷希尔（Melvin Drescher）发明了囚徒困境，[192]用来研究这种合作难题。

囚徒困境基本形式如下。两个人（姑且称他们为爱丽丝和鲍勃）因为合伙犯罪被逮捕了，警察将他们分别关在不同的房间（图14.2）。爱丽丝和鲍勃必须各自决定是否指证对方。如果爱丽丝同意指证鲍勃，她就会被释放，而鲍勃则会被判处无期徒刑。但如果爱丽丝拒绝指证，而鲍勃同意指证爱丽丝，则他将被释放，爱丽丝则会被判无期徒刑。如果两人都指证对方，则两人都会被关入监狱，但是只判10年徒刑。而如果两人都不指证对方，则两人都会判得较轻，只需入狱5年。警察要求他们不许相互沟通，必须马上做出决定。

如果你是爱丽丝，你会怎么做呢？

你可能会这样想：鲍勃有可能指证你，也有可能不指证，你不知道他会怎么做。如果他指证你，你最好的选择就是指证他（入狱10年要好于无期）。如果他不指证你，你最好的选择仍然是指证他（释放要好于入狱5年）。因此，不管鲍勃怎么做，你脱罪最好的选择都是同意指证他。

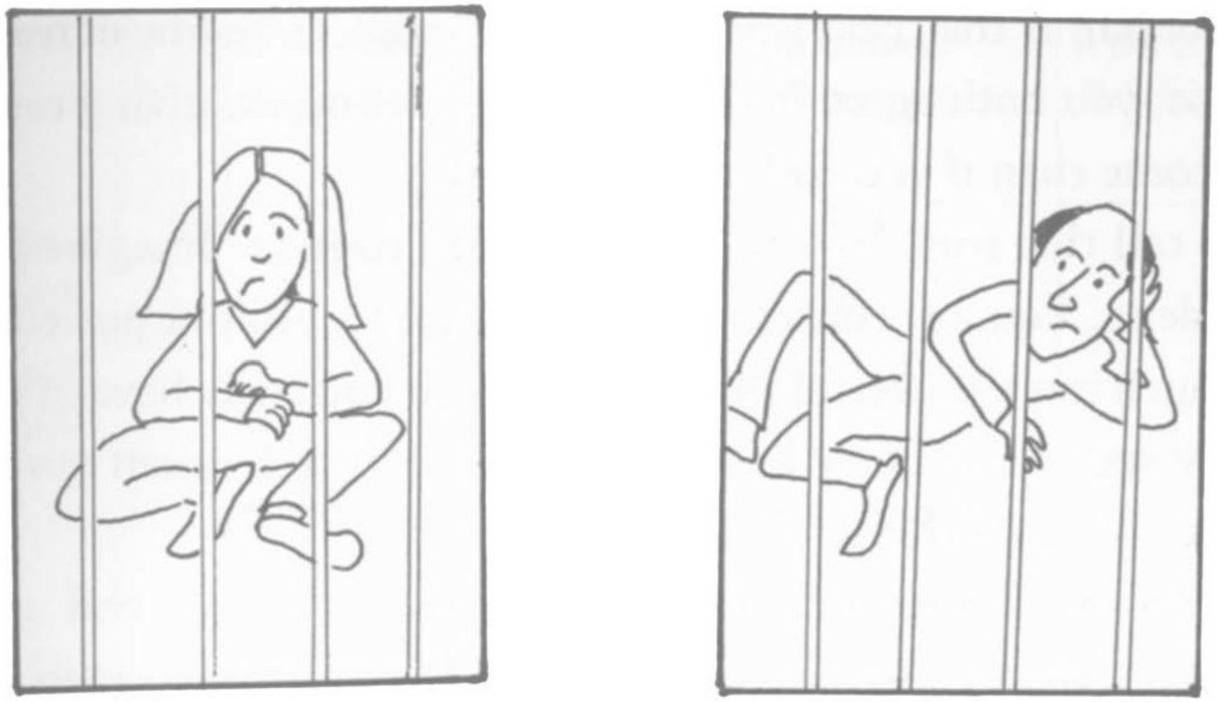


图14.2　爱丽丝和鲍勃面临着“囚徒困境”（David Moser绘制）

问题是鲍勃也会这样想。因此你们俩都会同意指证对方，而如果你们俩都保持沉默，得到的结果会更好。

我们再换一个场景来看这个问题，假设你是美国总统，你在考虑是否建造一种强大的核武器，比你现有的武器强得多。你怀疑俄国政府也在考虑这样做，但是不确切。

假设俄国决定发展这种武器，如果你也决定建造这种武器，那美国和俄国的军力就会维持平衡，尽管两国的花费都不少，世界也处于更危险的境地；如果你决定不发展这种武器，那么俄国对美国的军力就会占优势。

假设俄国决定不发展这种武器，如果你决定建造，那美国对俄国的军力就会占优，同时国家也会有财政负担；而如果你决定不建造，那么美国和俄国仍然会维持军力平衡。

同鲍勃和爱丽丝面临的情形一样，无论俄国怎么做，你最好的选择都是同意建造，因为不管俄国的选择是什么，建造武器都是美国更好的选择。当然，俄国也这样想，因此两国都会决定建造新的核弹，而如果

两国都不建造，本来结果会更好。

这就是囚徒困境悖论——用政治学家阿克塞尔罗德（Robert Axelrod）的话说，“每个人都追求自利，使得所有人的利益都受损。”[193] 这个悖论指的是群体中的个体由于只顾自身利益，整体上却使得群体所有个体都受损的情形（全球变暖就是典型的这种例子）。经济学家哈丁（Garrett Hardin）有个著名的词描述这种情况——“公地悲剧”。[194]

囚徒困境及其变体作为理想模型体现了合作问题的本质，其影响遍及学术界和商业界，并且改变了各国政府对于核裁军、恐怖主义以及合作管理和规范等现实世界中政策问题的看法。

囚徒困境通常用两人“博弈”的收益矩阵表示——矩阵中各元素为两个人在各种可能情形下的收益。表14.1给出了囚徒困境收益矩阵的一个例子。博弈的目标是尽可能多得分（蹲监狱的时间尽可能短）。参与者各自决定“合作还是背叛”，各决定一次形成一个回合。也就是说，每个回合中，参与者A和B不能相互商量是否合作（例如，拒绝指证；决定不建造核弹）。如果双方都合作，各得3分。如果A合作而B背叛，则A得0分而B得5分，反之则A得5分而B得0分。如果双方都背叛，则各得1分。前面说了，如果只进行一个回合，则两人合理的策略就是背叛。但如果有多个回合，也就是说，如果双方相互博弈多次，则总是背叛的参与者的收益会远低于学会了相互合作的参与者。互利合作是如何产生的呢？

表**14.1**　囚徒困境的收益矩阵

	A合作 B合作	A合作 B背叛	A背叛 B合作	A背叛 B背叛
A	3	0	5	1
B	3	5	0	1

密歇根大学的政治学家阿克塞尔罗德（图14.3）深入研究了囚徒困境。他在这方面的成果深深影响了许多学科，也让他赢得了许多奖项，

包括麦克阿瑟“天才”奖。

阿克塞尔罗德因为关注军备竞赛，从冷战时期开始研究囚徒困境。他的问题是：“在一个自私的世界里，如果没有中央权威，合作要如何才能出现？”[195] 阿克塞尔罗德注意到历史上对这个问题最著名的回答是17世纪哲学家霍布斯（Thomas Hobbes）给出的，他认为合作只有在存在中央权威的情况下才有可能产生。[196] 300年（和无数场战争）之后，爱因斯坦也提出了类似的观点，[197] 认为在核武器时代要确保和平的唯一途径就是成立高效的世界政府。国际联盟，以及后来的联合国，就是以此为目的成立的，但是都没有成功成立世界政府，或是维持世界和平。

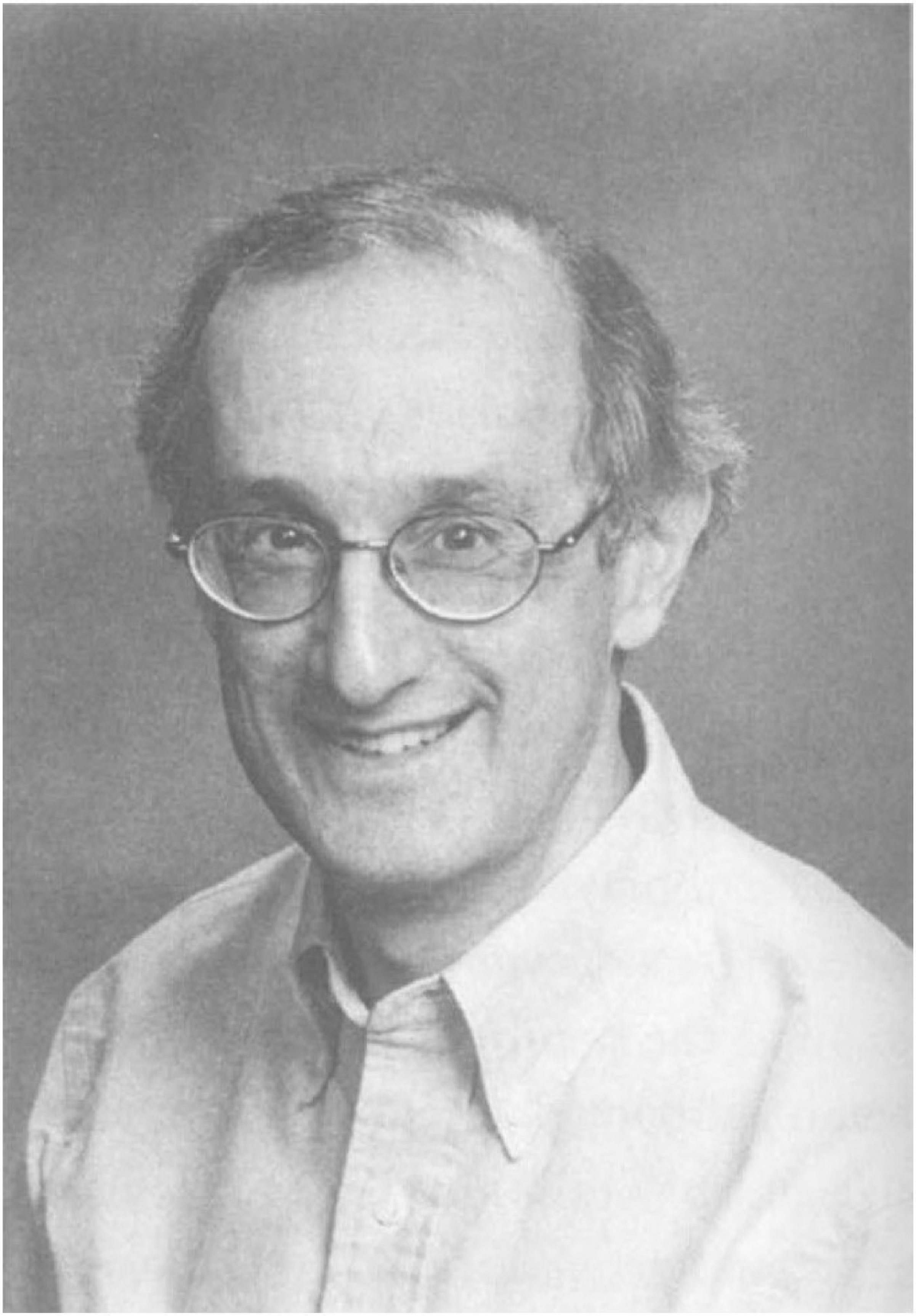


图14.3　阿克塞尔罗德（密歇根大学复杂系统研究中心提供照片）

高效的世界政府看来是指望不上了，阿克塞尔罗德想知道，如果没有世界政府是不是也有可能产生合作。他认为通过研究多回合的囚徒困境也许能有助于认识这个问题。阿克塞尔罗德认为，“合作要能够产生”就意味着，不管对手的策略如何变化，从长期来看，合作策略必须比非合作策略的收益更高。而且，如果用达尔文选择对策略进行演化，则种群中的合作策略的比重应当会随时间增加。

用计算机模拟囚徒困境

阿克塞尔罗德想知道什么样的策略是好策略，因此他组织了两次囚徒困境竞赛。他让各学科的研究人员提出各自的策略，并根据策略设计能进行囚徒困境博弈的计算机程序，然后在比赛中让这些程序互相博弈。

回想一下第9章讨论的机器人罗比，策略指的其实就是一组规则，规定了在各种情形中应该采取何种行动。对于囚徒困境来说，策略就是根据对手以前的行为决定下一步是合作还是背叛的规则。

第一次竞赛收到了14个程序；第二次增加到了63个。每个程序都相互博弈200个回合，根据表14.1中的收益矩阵计算总分。程序可以有记忆——每个程序都能存储一些之前与对手的博弈记录。有些提交来的策略相当复杂，使用统计方法分析其他策略的“心理”。然而，两次竞赛中获胜的策略——平均得分最高——都是所提交的策略中最简单的策略：针锋相对（TIT FOR TAT）。这个策略是数学家拉普波特（Anatol Rapoport）提交的，第一个回合合作，然后在后面的回合中采取对手在前一回合中所使用的策略。也就是说针锋相对策略愿意合作，并且对愿意合作的对手以礼相待。但如果对方背叛，针锋相对策略就会回之以背叛，直到对手又开始合作为止。

让人吃惊的是，这样简单的策略竟然打败了其他所有策略，尤其是第2次竞赛时参赛者已经知道了针锋相对策略，可以有针对性地与它对抗。然而，在参赛的几十位专家中，没有人能设计出更好的策略。

根据竞赛结果，阿克塞尔罗德得出了一些一般性结论。他注意到所有成绩好的策略都有友善的特点——他们从不先背叛对方。友善的策略

中得分最低的是“绝不宽恕”策略：它开始时合作，但一旦对方背叛，它以后就会一直背叛。而针锋相对策略会以一次背叛惩罚对手的上一次背叛，但如果对手又开始合作，它就会原谅对方，也开始合作。阿克塞尔罗德还注意到，虽然大多数成功的策略既友善也能宽恕对手，但是它们也具有报复性——它们会在背叛发生后很快就进行惩罚。针锋相对策略不仅友善、宽恕和进行报复，它还有一个很重要的特性：行为明确，具有可预见性。对手很容易就能知道针锋相对采取的策略，因此也就能预知它会如何对对手的行为做出反应。这种可预见性对于促进合作很重要。

有趣的是，阿克塞尔罗德在竞赛之后又进行了一系列实验，用遗传算法演化囚徒困境的策略。策略的适应度就是它与种群中其他策略反复博弈之后的得分。遗传算法演化出的策略行为与针锋相对也是一样的。

囚徒困境的扩展

阿克塞尔罗德对囚徒困境的研究在20世纪80年代引起了轰动，尤其是在社会科学中影响很大。人们开始研究它的各种变体——采用不同的收益矩阵、不同的参与人数或者在多方博弈中让各方能选择对手，等等。其中有两个很有趣的实验分别增加了社会规范和空间结构。

加入社会规范

阿克塞尔罗德进行了添加社会规范的囚徒困境实验，[198] 实验中规范表现为在背叛被其他人发现时的社会谴责（用负分表示）。在阿克塞尔罗德的多方博弈中，个体的每次背叛，都有一定的概率被其他个体发现。除了决定合作或背叛的策略，每个个体还有在发现其他个体的背叛行为时决定是否进行惩罚（减分）的策略。

具体说，每个个体的策略由两个数字组成：背叛的概率（无耻度）和发现背叛行为时进行惩罚的概率（正义度）。在最初的群体中，概率值被随机赋予。

在每一代，群体进行一次循环博弈：群体中每个个体与其他所有个体博弈一次，每当出现背叛，背叛行为都有一定概率被其他个体发现。一旦被发现，发现背叛行为的个体就会根据自身的正义度以一定概率对背叛个体进行惩罚。

在每次循环之后，会发生进化过程：根据适应度（得分）选择父代策略，从而产生出下一代策略。父代通过变异复制产生后代：每个后代的无耻度和嫉恶度在父代的基础上稍微变化。如果开始时群体中大多数个体的正义度都设为0（也就是没有社会规范），那背叛者就会越来越多。阿克塞尔罗德最初希望能发现促进群体中合作进化的规范——也就是说，进化出正义度以对抗无耻度。

然而，结果是仅仅有规范并不足以保证产生合作。在后来的实验中，阿克塞尔罗德又加入了元规范（metanorms），在其中有执法者来惩罚非执法者。不知道你们明不明白我的意思，就好像逛超市的时候，如果我没有阻止我的小孩在过道里嬉闹，还撞到了其他顾客，有些这样的人就会以鄙视的眼神看着我。这样的元规范对我很有效。阿克塞尔罗德也发现元规范很有用——如果周围有惩罚者，非惩罚者就会演化得更倾向于惩罚，而被惩罚的背叛者也会演化得更愿意合作。用阿克塞尔罗德的话说，“元规范能促进并保持群体中的合作”。[199]

加入空间结构

数学生物学家诺瓦克（Martin Nowak）和其合作者在囚徒困境中加入了空间结构，这种扩展也非常有趣。在阿克塞尔罗德最初的实验中没有空间的概念——所有参与者相遇的可能性都一样，参与者之间的距离没有意义。

诺瓦克猜测，如果参与者位于空间网格上，有严格的邻居概念，则会对合作的进化产生很强的影响。诺瓦克同他的博士后导师梅一起进行了计算机仿真[200]（在第2章讲逻辑斯蒂映射时我们已经见过梅），实验中参与者位于2维网格上，都只与最近的邻居博弈。如图14.4所示，图中有一个5×5的网格，每个格子中有一个参与者（诺瓦克和梅的网格要大得多）。所有参与者的策略都极为简单——它们没有记忆；要么一直合作，要么一直背叛。

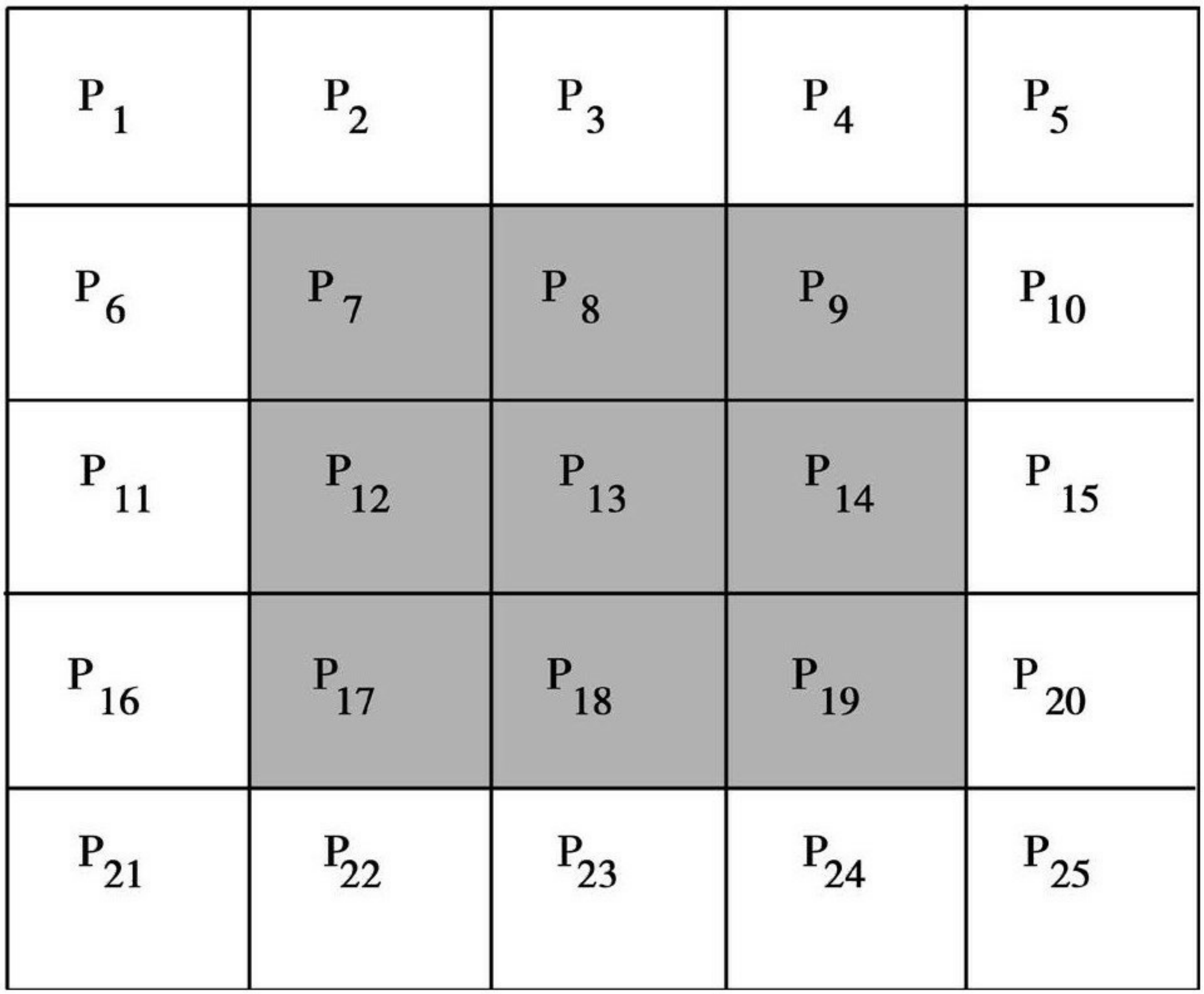


图14.4　空间囚徒困境博弈示意图。每个策略只与最近的邻居进行博弈——例如，策略P13只与其最近的邻居（阴影部分）进行博弈和生存竞争

模型周期运行。在每个时间步，每个策略与8个最近的邻居分别进行一次囚徒困境博弈（与元胞自动机类似，边缘上的格子回绕相连），然后将8次博弈的得分加总。接下来进行选择，每个策略都替换为邻域中得分最高的策略（有可能是它自己）：没有变异。

这项研究的出发点是生物学。就像诺瓦克和梅说的，“我们认为，群体中确定的空间结构对合作的进化可能很重要[201]，无论是分子、细胞，还是生物”。

诺瓦克和梅用各种合作和背叛个体的初始设置以及不同的收益矩阵进行了实验。他们发现在有些条件下，合作和进化个体的空间分布模式有可能振荡，甚至产生“混沌性变化”，[202]　在其中合作者和背叛者共

存。而在非空间性的多方囚徒困境博弈中，如果没有前面提到的元规范，最后背叛者会占据群体。在诺瓦克和梅加入空间后，合作者可以一直坚持下去，无需在博弈中加入规范或元规范。

诺瓦克和梅认为他们的结果表现了真实世界中的一个特性——即，空间相邻关系的存在会促进合作。在评论这项研究时，生物学家西格蒙德（Karl Sigmund）这样说道："地域性有利于合作 [203] ……这在真实的社区中可能也是成立的。"

建模的好处

对于囚徒困境这样的理想模型的计算机仿真，如果做得好的话，会是实验科学和数学理论的有力补充。对于大部分复杂系统来说，不可能对其进行真正的实验，用数学研究也非常困难，这个时候模型就是研究它们的唯一可行途径。像囚徒困境这样的理想模型最大的贡献就是为我们提供了一把钥匙，来研究合作的进化这类没有精确的科学术语和完善定义的现象。

囚徒困境模型具备了前面列出的科学中理想模型的所有作用（其他复杂系统模型也可以列出类似的贡献）：

证明了解释某种现象的机制是不是合理。例如，霍布斯可能不会相信，在自私但具有适应性的个体组成的缺乏领袖的群体中会产生合作，而各种各样的囚徒困境模型则证明了这一点（虽然是以某种理想形式）。

研究简单模型在改变后的效应，引导对复杂现象的认识。各种各样的囚徒困境变体揭示了合作得以产生的条件。比如说，你可能会问，如果本来想合作的人犯了错，不小心发出了不合作的信号，打个比方，美国总统的意见被翻译成俄语时不小心弄错了，会发生什么呢？囚徒困境模型为研究这类讯息错误的影响提供了条件。霍兰德曾把这种模型比作"飞行模拟器"[204]，可以用来验证人们的思想和改进人们的认识。

为新技术带来灵感。囚徒困境的研究结果——例如，合作得以产生和维持的条件——就被用来帮助改进P2P网络 [205] 和在电子商务中防止欺诈 [206]，类似的应用还有很多。

引出数学理论。一些人利用囚徒困境的计算机仿真结果研究合作产生条件的通用数学理论。最近的一个例子就是诺瓦克在题为《合作进化的5条规则》的文章中所做的工作。[207]

从囚徒困境这样的理想模型中得出的结论能否用于改进政府外交策略或全球变暖等问题的政策决策呢？用理想模型研究各种政策效果的想法很有吸引力，事实上，囚徒困境及其相关模型对政策分析的影响已经很大了。

举个例子，新能源金融（New Energy Finance）是专为全球变暖问题提供咨询的公司，这家公司最近公布了一份报告，[208]题为《如何拯救地球：友善、报复、宽恕、明确》。报告中认为气候变化问题就是多方多回合的囚徒困境问题，各个国家可以选择合作（减少碳排放，代价是经济受损）或是不合作（什么也不做，目前来看省了钱）。这个博弈会反复进行，不断缔结新的规范碳排放的协议和公约。报告认为，各国政府和国际组织的政策应当采取“友善、报复、宽恕、明确”的原则，正是阿克塞尔罗德在重复囚徒困境中指出的成功所需的条件。

另外，前面介绍了，通过研究规范和元规范的模型，发现规范和元规范对维持合作都很重要。这个结果对政府制定应对恐怖主义、军备控制和环境治理等问题的政策也有影响。[209]诺瓦克和梅的空间囚徒困境模型则告诉人们，在许多问题中——无论是生物多样性的维持和细菌制造新抗生素的效率问题[210]——空间和位置对于促进合作起到的作用。

计算机建模注意事项

> 所有模型都是错的，但是有一些有用。[211]
>
> ——伯克斯（George Box）和德雷珀（Norman Draper）

前面描述的模型都很简单，但是却有力地促进了科学的进步和政策的调整。它们带来了全新的认识，为研究复杂系统提供了新的途径，也让我们对怎样建立有用的模型有了更深刻的理解。不过，对于这些模型结果的意义以及它们在真实世界中的应用，却存在着一些夸大其词。因此，科学家应当仔细审视这些模型，弄清楚所得到的结果的普遍性。最好的方法就是看看这些结果是不是可重复。

在天文学或化学这样的实验科学中，所有重要的实验都必须是可重复的，也就是说其他科学家如果进行同样的实验要能得出同样的结果。如果实验结果没有人能重复，就不会（也不应当）被采信。不计其数的科学观点都是因为他人无法重复而被判了死刑。

计算机模型也必须是可重复的——也就是说，其他人重新构造所提出的模型要能得到同样的结果。阿克塞尔罗德就极力拥护这种观点，他写道："可重复性是科学积累的基石。 [212] 必须确认得到的仿真结果是否可靠，也就是说可以从头进行复制。如果没有进行确认，所发表的结论中有些就有可能只不过是程序错误导致的，歪曲了所仿真的对象，或是分析结果存在错误所致。可重复性对于检验模型结论的稳健性也很有用。"

大部分研究者从内心接受这样的观点，尝试重复了许多著名的囚徒困境仿真。这其中也产生了一些有趣甚至出人意料的结果。

1995年，胡伯尔曼（Bernardo Huberman）和格兰斯（Natalie Glance）重复了诺瓦克和梅的空间囚徒困境模型。[213] 他们的仿真只有一处改动。在原来的模型中，每一步格子上所有策略的博弈都同时进行，然后再在邻居中选择适应度最高的策略。（诺瓦克和梅必须在非并行计算机上模拟这种并行性。）胡伯尔曼和格兰斯则让一些博弈可以异步进行——也就是说，一些策略先进行博弈并选择，然后另一些策略再接着做，这样轮着来。这样简单的变化，应该说是让模型更符合实际了，他们却发现结果经常是整个网格中合作者完全被不合作者取代。穆科吉（Arijit Mukherji）、拉詹（Vijay Rajan）和斯莱格勒（James Slagle）也独立得到了类似的结果。[214] 他们还发现，只要存在一点失误或是欺骗（例如，合作者无意或是有意地背叛），合作就无法继续。诺瓦克、梅和他们的合作者博恩霍艾弗（Sebastian Bonhoeffor）则回应说，[215] 这样的改变只有在收益矩阵取一定的值时才会导致合作者消失，而在其他情况下则不会，至少在很长的时间内都不会消失。

2005年，加兰（Jose Manuel Galan）和利兹奎尔多（Luis Izquierdo）发表了他们重复阿克塞尔罗德的规范和元规范模型实验的结果。[216] 在阿克塞尔罗德的实验后已经过去了20年，计算机的性能已得到大幅提升，他们可以让仿真运行更多的周期，也可以彻底研究如果改变收益矩阵、变异概率等条件会导致什么结果。他们的结果与阿克塞尔

罗德的一些结果相吻合，但也有一些结果相当不同。例如，他们发现虽然元规范在短期内会促进和维持合作，但如果仿真很长时间，不合作者最终还是会占据优势。他们还发现结果对收益矩阵等条件相当敏感。

我们应当怎样理解这一切呢？我认为就像伯克斯和德雷珀说的：所有模型都是错的，但是有一些对于尝试研究极为复杂的系统却很有用。独立的重复实验能够发现理想模型中隐藏的一些不切实际的假设和对某些参数的敏感性。当然重复实验本身也应当被重复检验，就像实验科学一样。最后，建模者也必须明确指出模型的局限性，以免模型的结果被误读，生搬硬套，或是过分渲染。我用囚徒困境的模型作为例子来说明这些观点，其他复杂系统的简化模型同样应当注意这些。

最后借用一段物理学家安德森（Phillip Anderson）在1977年诺贝尔奖授奖仪式上讲的一段话（他也是建模的先驱）：

> 建模的艺术就是去除实在中与问题无关的部分，[217] 建模者和使用者都面临一定的风险。建模者有可能会遗漏至关重要的因素；使用者则有可能无视模型只是概略性的，意在揭示某种可能性，而太过生硬地理解和使用实验或计算的具体结果样本。

4　网络

在爱希莉亚[218]，城的生命是靠各种关系维持的。为了建立这些关系，它的居民从房子的角落拉起绳子，或白或黑或黑白相间，视关系的性质——血缘、贸易、权力、代表——而定。绳子愈来愈多，到了走路都通不过的时候，居民就会离开：只留下绳子和系绳子的东西。

带着财产露宿的爱希莉亚难民，从山边回望平原上那竖起木柱和绷紧绳索的迷宫，它仍然是爱希莉亚城，而他们则不算什么。

他们在另一个地方再建爱希莉亚。他们织起另一张类似的绳网，希望它比以前那一张更精细更有规律。后来他们又放弃了，把房子搬到更远的地方。

因此，在爱希莉亚境内旅行的时候，你会看到一些被舍弃的城的废墟，不耐用的墙已经失踪了，死去的骸骨也被风卷走了：一些纠缠不清的关系的蛛网在寻找形式。

——卡尔维诺（Italo Calvino），

《看不见的城市》（Invisible Cities ）

第15章　网络科学[219]

小世界

我住在俄勒冈州波特兰，这个市区大约有200万人。我在波特兰州立大学（Portland State University，PSU）任教，学校有将近25000名学生，超过1200名教师。几年前，我们家换了新房子，离学校较远。有一次我同我们的新邻居桃乐茜聊天，她是位律师。我告诉她我在波特兰州立大学教书。她说：“不知道你认不认识我父亲。他叫乔治·勒恩达理斯（George Lendaris）。”我很吃惊。勒恩达理斯是我在PSU的同事，整个学校只有三四个老师研究人工智能，其中就包括我们俩。就在前天，我

还和他见了面，讨论合作申请经费。这世界真小！

几乎所有人都有过这种“小世界”经历，很多比我遇到的更具戏剧性。我丈夫高中最好的朋友和我在人工智能课上采用的课本的作者是堂兄弟。在圣塔菲住在离我三栋房子里的一位女士是我在洛杉矶的高中英语老师的好友。我相信你也有过类似的经历。

这种出人意料的关系到底有多常见呢？20世纪50年代，哈佛大学的心理学家米尔格兰姆（Stanley Milgram，图15.1）对这个问题产生了兴趣，他想弄清在美国一个人平均要通过几个熟人关系才能到达另一个人。他设计了一个实验，实验中一些普通人被要求将一封信寄给一位陌生人，他可以将信交给他认为最有可能将信送达的熟人，熟人又转交给熟人的熟人，直到信通过熟人关系形成的链条送到收信人手中。

米尔格兰姆在报纸上刊登广告，在堪萨斯州和内布拉斯加州招募了一群“发信人”，告诉他们“收信人”的姓名、职位和所在城市，发信人要把信送给他不认识的这位收信人。米尔格兰姆选择的收信人中，有一个例子是波士顿的一位股票经纪人，还有一个例子是坎布里奇（Cambridge）附近一位神学学生的妻子。发信人被要求将信送给他认识的某位熟人，再请这位熟人继续传送。传送过程被记录在信上，如果信送到了收信人手里，米尔格兰姆就计算信经过了几个熟人关系。米尔格兰姆记述了一个例子[220]：

> 在信封被交给堪萨斯州一位发信人4天后，圣公会神学院的一位教师在街上拦住了我们的收信人。他将一个牛皮信封塞给她：“爱丽丝，这是你的。”一开始她以为这是一封没有送到发信人手里被退回来的信，从没有离开过坎布里奇，但是当我们看上面的记录时，我们惊喜地发现信是堪萨斯州的一位农夫寄来的。他将信交给了他们当地圣公会的牧师，这位牧师又将信寄给了在坎布里奇任教的这位牧师，这位牧师再将信交给了收信人。这样从发信人经过两个熟人关系就到了收信人！



图15.1　米尔格兰姆（1933—1984）（Eric Kroll摄影，经亚历山德拉·米尔格兰姆夫人许可使用）

在这项著名的实验中，米尔格兰姆发现，在送达的信件中，从发信人平均经过5个熟人就送到了收信人的手中。这个发现后来广为人知，被称为“六度分隔（six degrees of separation）”。

后来心理学家柯兰菲尔德（Judith Kleinfeld）研究发现，[221] 米尔格兰姆的发现被曲解了——事实上，大部分信件从没有到达收信人手中，而在米尔格兰姆的其他研究中，到达收信人的信件经过的平均熟人关系也不止5个。然而，六度分隔的小世界思想还是成了我们文化的传奇[222]。正如柯兰菲尔德指出的：

> 当人们发现出人意料的社会关系时[223]，很有可能会印象深刻……在理解自然界的巧合时，我们的数学水平不高，直觉也不咋样。

那这到底是不是一个小世界呢？这个问题最近又受到很多关注，不仅仅是社会网络，还涉及其他各种网络，包括活细胞中的代谢网络和遗传调节网络，以及增长迅猛的万维网。过去十年中，这些网络的问题吸引了无数复杂系统研究者，从而产生了所谓的“网络新科学”[224]。

网络新科学

你肯定看到过类似于图15.2这样的网络图。这是大陆航空（Continental Airlines）在美国的航线图。点（或节点）代表城市，线（或连接）代表城市之间的航班。

许多自然、技术和文化现象经常被描述为网络，航线图就是一个明显的例子。大脑是神经元通过突触连接起来的巨大网络。细胞中的遗传活动是受由基因通过调节蛋白质连接起来的复杂网络控制。社会则是由各种各样的关系连接起来的人（或组织）组成的网络。万维网则更是现代社会的典型网络。在国家安全领域，识别和分析可能的“恐怖分子网络”是很重要的工作。直到不久前，网络科学都不被视为一个研究领域。数学家研究抽象网络结构的学科被称为“图论”，神经科学家研究神经网络，流行病学家研究疾病通过人际网络的传播。像米尔格兰姆这样的社会学家和社会心理学家则研究社会网络的结构。经济学家研究经济

网络的行为，例如技术革新在商业网络中的传播。航空公司主管则研究图15.2这样的网络，想找到在一定条件下能获得更多利润的网络结构。他们基本上都是各干各的，通常都互相不知道其他人的工作。

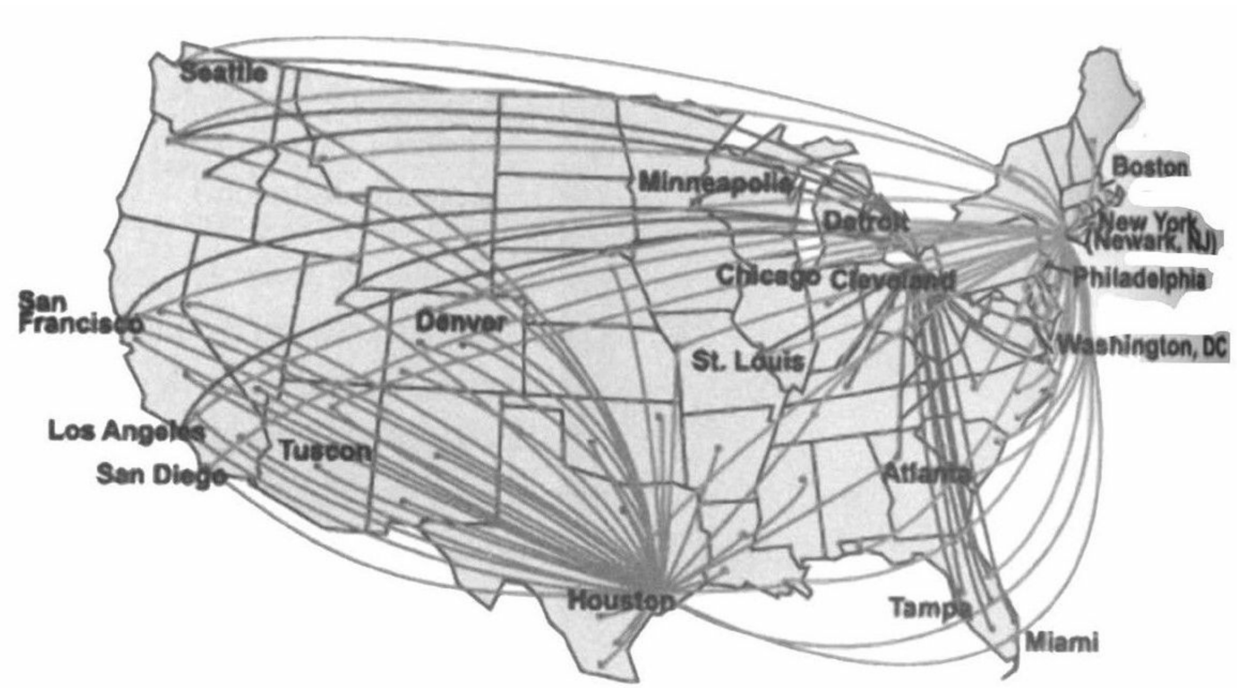


图15.2　大陆航空公司航线简图（图片来自NASA虚拟天空；http：//virtualskies.arc.nasa.gov/research/tutorial/tutorial2b.html）

不过，近十年来，越来越多的应用数学家和物理学家开始着迷于研究一系列操控所有自然、社会和技术网络的普适原理。这股网络研究浪潮是由20世纪90年代末的两篇重要文章[225]引发的：邓肯·瓦特（Duncan Watts，图15.3）和斯托加茨（Steven Strogatz，图15.4）的《‘小世界网络’的集体动力学》，以及巴拉巴西（Albert-László Barabási，图15.5）和艾伯特（Réka Albert）的《随机网络中标度的涌现》。这两篇文章分别发表在世界上最著名的科学期刊《自然》和《科学》上，很快就引起了巨大反响。关于网络的各种新发现迅速涌现。

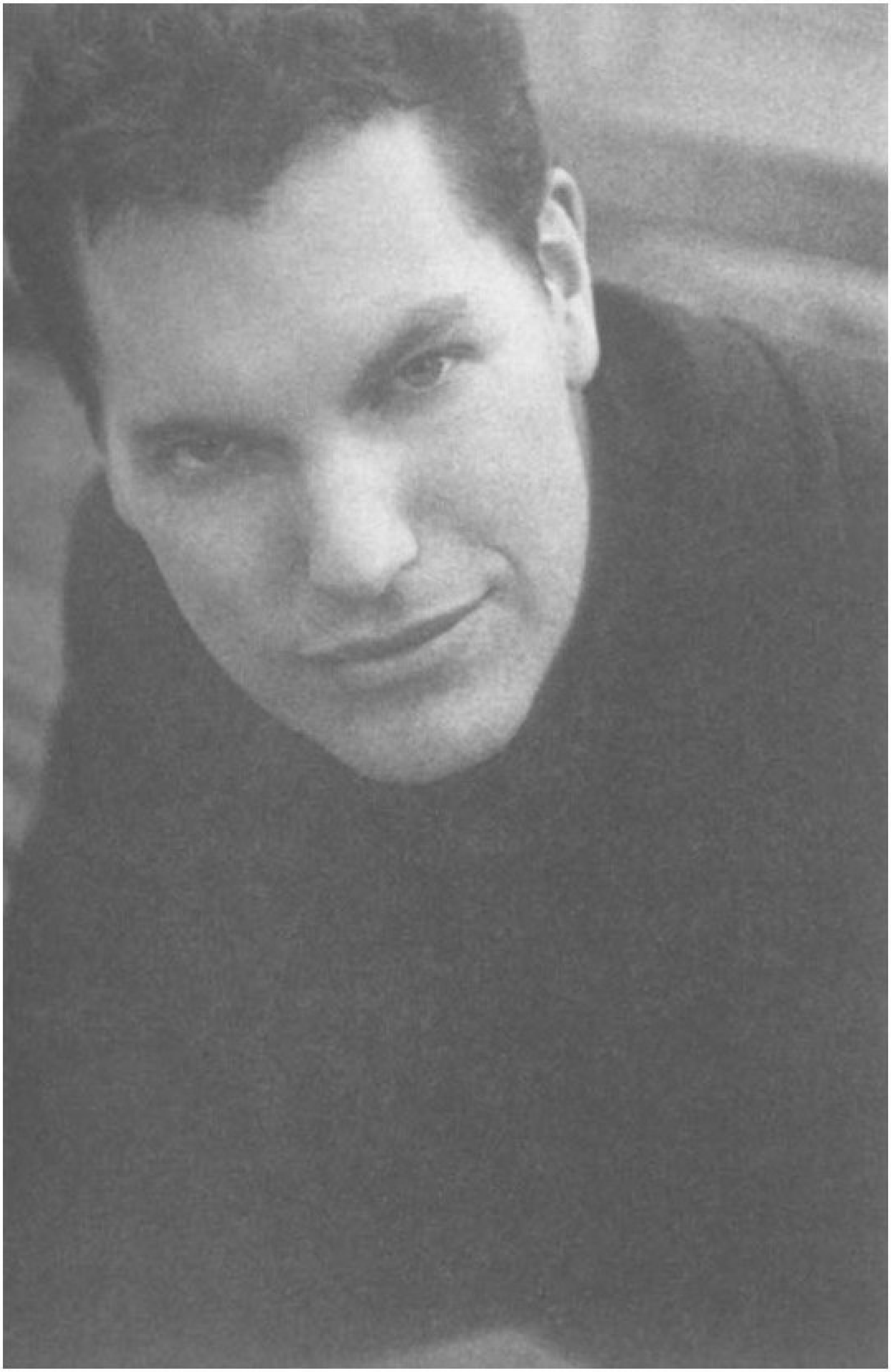


图15.3　邓肯·瓦特（经邓肯·瓦特许可使用）

网络科学的兴起恰逢其时。对各种网络的共性的研究，无论是仿真还是统计大量真实数据，都只有在计算机的速度足够快之后才可能做到。在20世纪90年代，条件成熟了。不仅如此，随着万维网在社会、商业和科学网络中越来越普及，大量数据都能很容易地得到。

此外，许多非常聪明的物理学家已经厌倦了越来越抽象的现代物理学，想要研究点别的东西。网络既包含数学和复杂系统动力学，又与现实世界相关，因此成为理想的研究对象。就像邓肯·瓦特说的（他自己是应用数学家和社会学家），“一大群饥肠辘辘的物理学家[226]兴奋地循着新问题的香味涌来”。

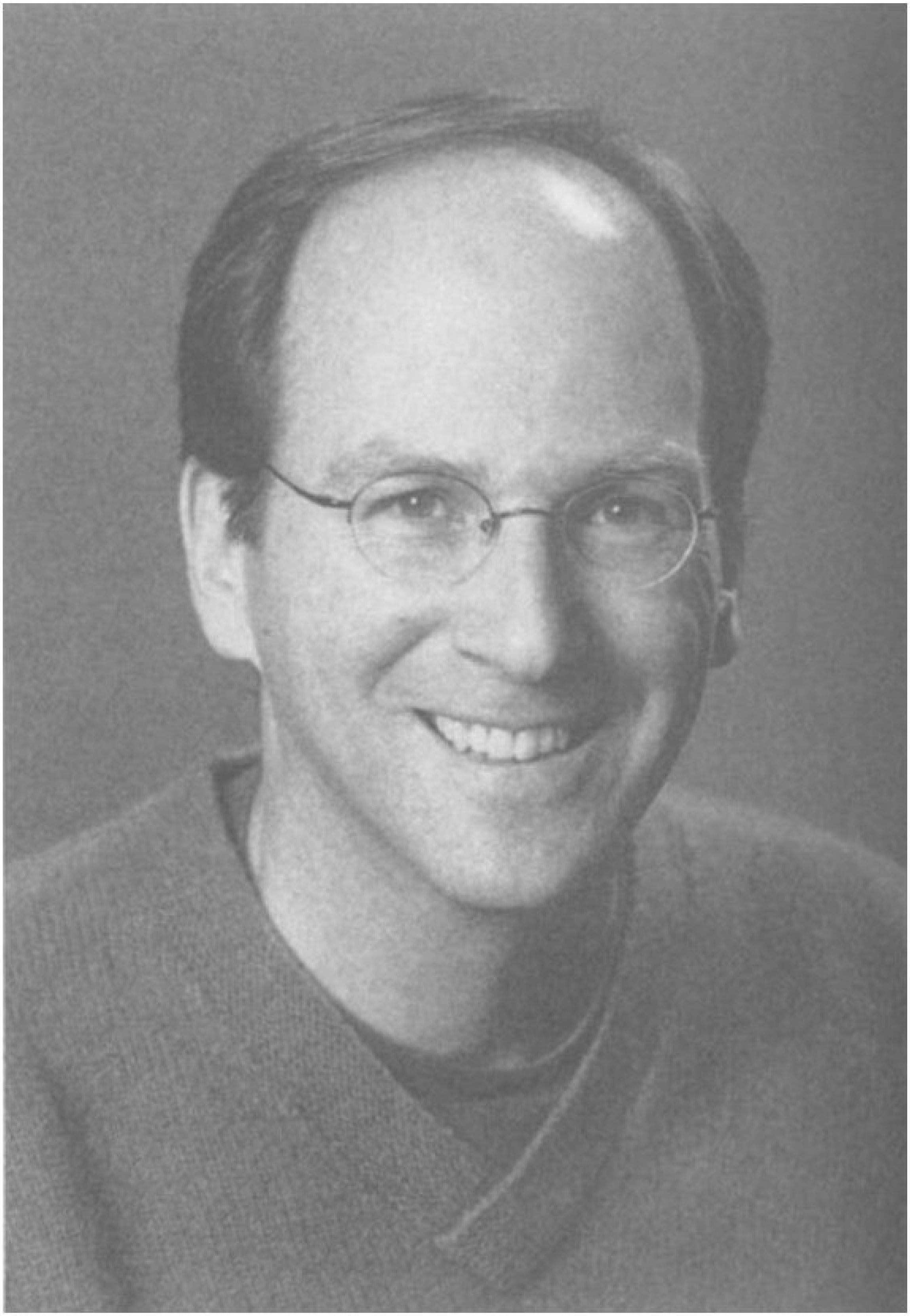


图15.4　斯托加茨（经斯托加茨许可使用）

这些聪明人掌握了合适的数学工具，能够将复杂问题简单化，同时

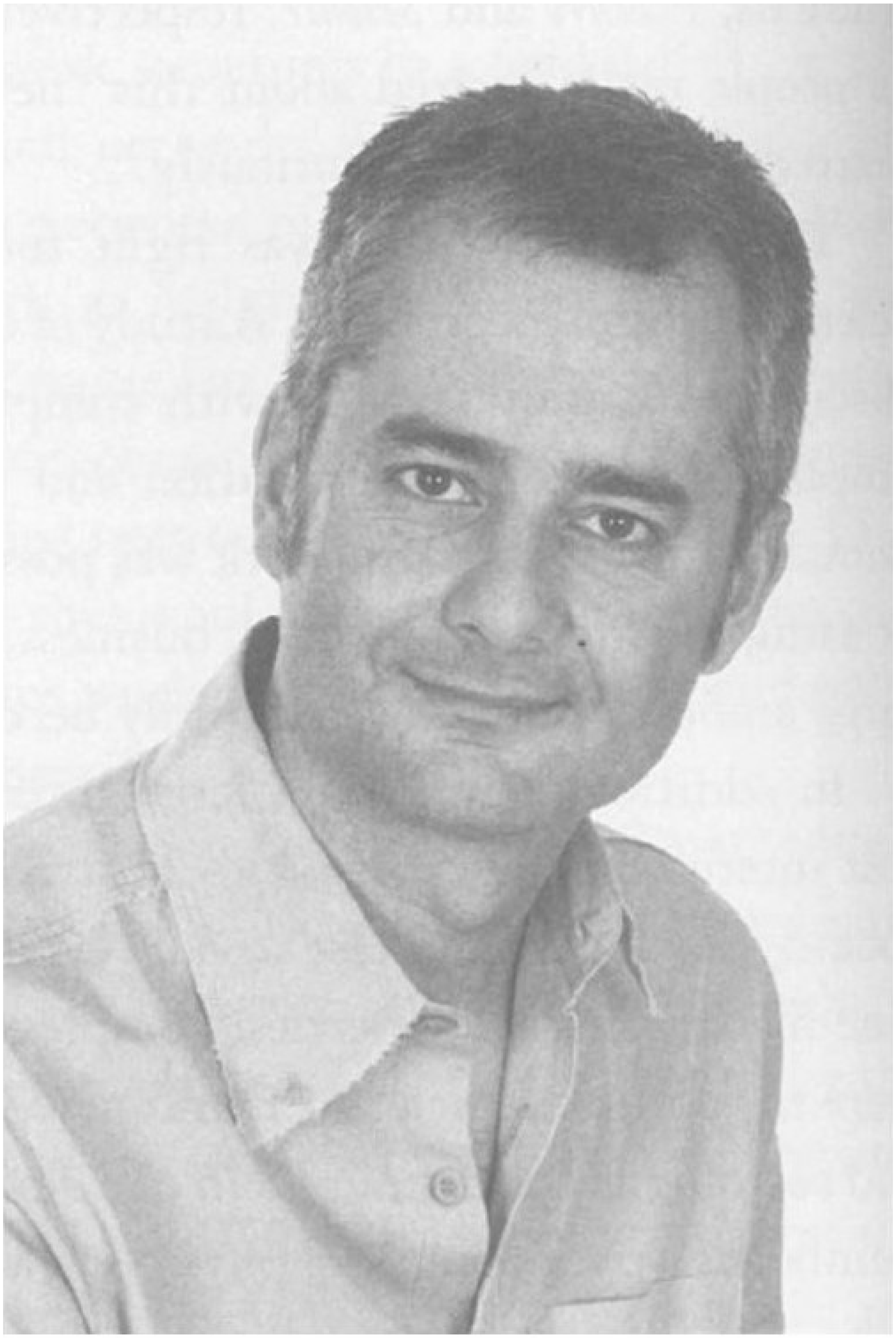


图15.5　巴拉巴西（经巴拉巴西许可使用）

又不丢掉本质特征。一些从物理学家转型而来的网络科学家已经成为这

个领域的领导者。

也许最重要的是，这些科学家逐渐意识到，各种高度复杂的网络系统对人类生活和福祉的影响越来越大，迫切需要有新的思想和方法——真正全新的思考方式——来帮助理解它们。

巴拉巴西将这种新方法称为“网络思维”，并认为“网络思维将渗透到人类活动和人类思想的一切领域”[227]。

什么是网络思维

网络思维意味着关注的不是事物本身，而是事物之间的关系。例如，第7章曾讲过，人类和芥菜都大致有25000个基因，这似乎无法体现人类与这种植物的生物复杂性的差别。近几十年来，一些生物学家提出生物的复杂性主要来自基因之间交互作用的复杂性。在第18章我们会详细讨论这一点，现在知道网络思维的最新成果对生物学有深刻影响就够了。

最近，网络思维还帮助厘清了一些看似无关的科学和技术之谜：为什么生物的生命期与它们的大小基本上遵循一个简单的函数？为什么谣言和笑话传播得如此之快？为什么电网和万维网这样大规模的复杂网络有时候非常稳健，有时候却又容易出现大范围崩溃？什么样的事件会让本来很稳定的生态群落崩溃？

这些问题看似毫不相干，网络科学家却认为答案反映了各种网络的共性。网络科学的目的就是提炼出这些共性，并以它们为基础，用共同的语言来刻画各种不同的网络。同时网络科学家也希望能理解自然界中的网络是如何发展而来的，以及它们是如何随时间变化的。对网络的科学理解不仅会改变我们对各种自然和社会系统的理解，同时也会帮助我们更好地规划和更有效地利用复杂网络，包括更好的网络搜索和万维网路由算法，控制疾病传播和有组织犯罪，以及保护生态环境。

到底什么是“网络”

要科学地研究网络，我们必须精确地定义网络的意义。用最简单的话说，网络是由边连接在一起的节点组成的集合。节点对应网络中的个

体（例如神经元、网站、人），边则是个体之间的关联（例如突触、网页超链接、社会关系）。

图15.6展示了一部分我自己的社会网络——一些我最密切的朋友以及他们的一些最密切的朋友，总共19个节点。（当然大部分“真实的”网络比这个要大得多。）初看上去，这个网络就像一团乱麻。然而，如果你仔细看一下，就会在这一团乱麻中发现一些结构。有一些联系紧密的群体——这不奇怪，我的一些朋友相互之间也是朋友。例如，David、Greg、Doug和Bob相互连接，Steph、Ginger和Doyne也是这样，我自己则是这两个群体之间的桥梁。不了解我的历史的人可能也猜得出这两个朋友“群体”与我的不同兴趣或不同的人生阶段有关。（两个答案都正确。）

你可能还注意到一些人有很多朋友（例如我自己、Doyne、David、Doug、Greg），一些人则只有一个朋友（例如Kim、Jacques、Xiao）。这是因为这个网络不完整，但是在大型社会网络中，总是有一些人有许多朋友，有一些人则朋友较少。

网络科学家创造了一些术语来描述各种类型的网络结构。网络中存在的内部联系紧密、外部较松散的群体被称为集群（clustering）。进出一个节点的边的数量称为这个节点的度（degree）。例如，我的度就是10，是所有节点中最高的。Kim的度为1，与其他5个人一起是最低的。借助这些术语，我们可以说网络中有少数高连接度的节点，以及大量低连接度的节点。

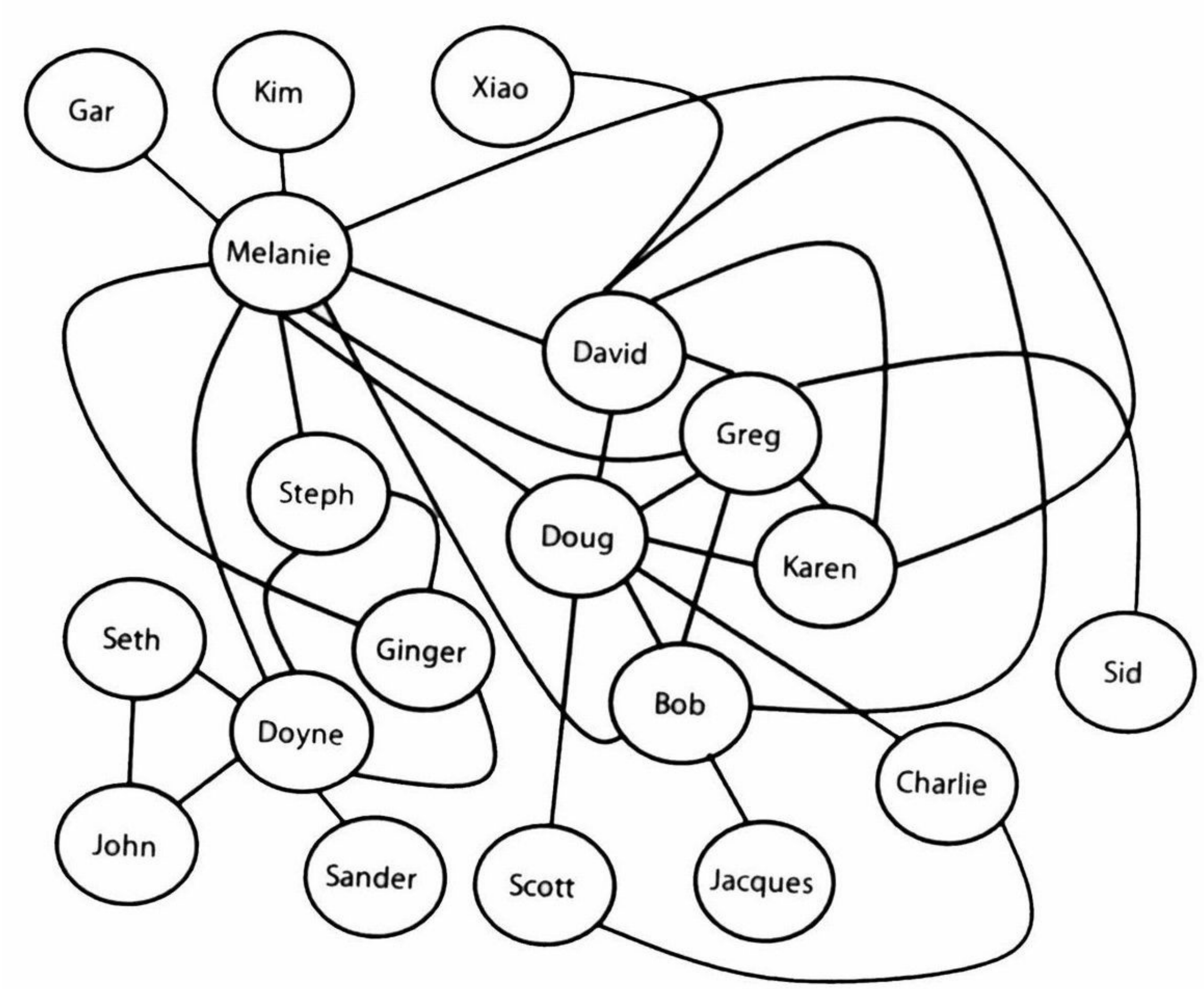


图15.6　一部分我自己的社会网络

在图15.7的网络度分布中，这一点可以看得很清楚。横坐标为度，长条的高度则对应具有这个度的节点的数量。例如，有6个节点的度为1（第一个长条），有1个节点的度为10（最后一个长条）。

从图中可以清楚看到大量节点具有低连接度，少量节点具有高连接度。在社会网络中，这表明大部分人的朋友相对较少，极少的人具有很多很多朋友。类似的，在万维网上，少数网站极受欢迎（很多网站都有链接指向这些网站），例如有超过7500万个链接指向谷歌，而大部分网站则几乎没什么知名度——例如只有123个链接指向我自己的网站[228]（其中大部分可能都来自搜索引擎）。

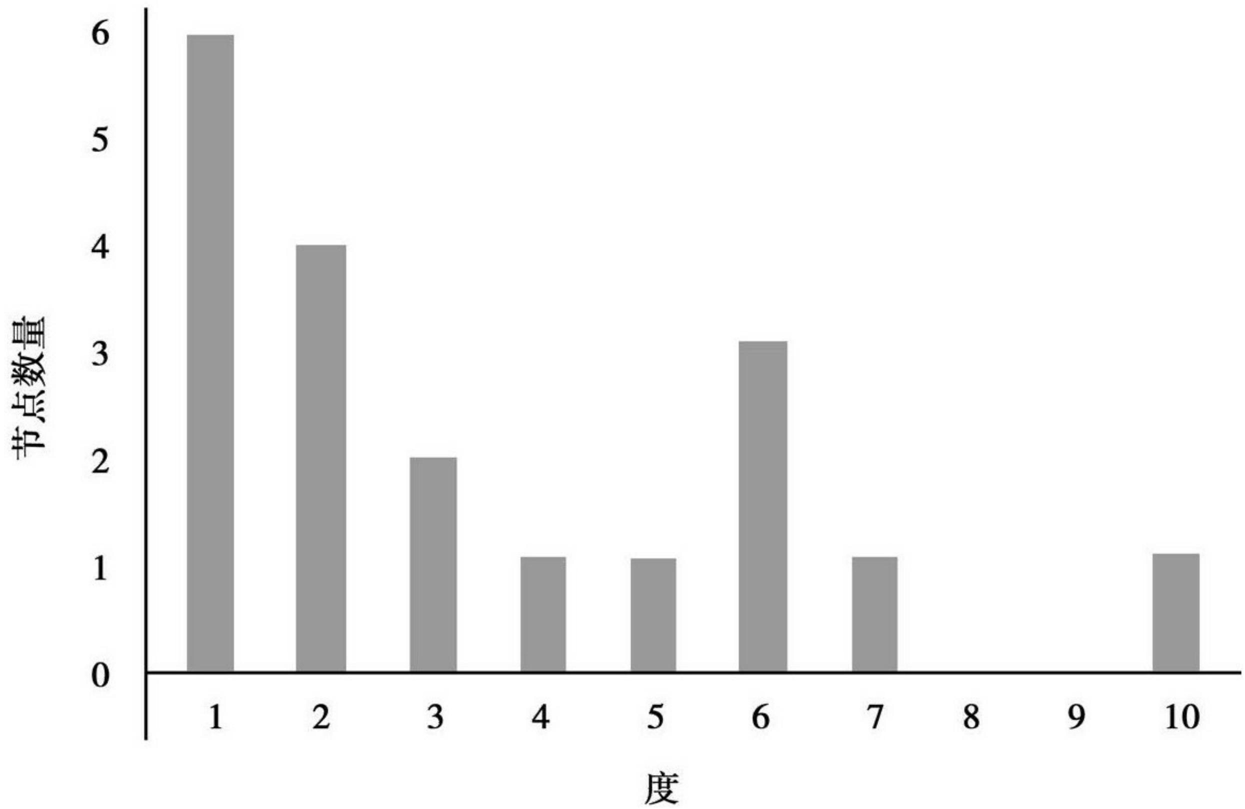


图15.7　图15.6中的网络的度分布。图中的长条代表具有相应度数的节点的数量

高连接度的节点被称为中心节点（hub），它们是网络中主要的信息或行为的传递渠道。图15.2显示了大多数航空公司在20世纪80年代解除管制后采用的中心节点系统：各家航空公司都选择特定的一些城市作为中心节点，大部分航班都经过这些城市。如果你曾坐大陆航空的航班从美国西部飞到东海岸去，你可能就会在休斯敦转机。

网络科学家发现，他们研究过的自然、社会和技术网络中，大部分都具有这些特征：高度的集群性、不均衡的度分布以及中心节点结构。这些特征的出现显然不是偶然的。如果我将节点随机连接起来生成一个网络，则所有节点的度数都会差不多，得到的度分布就不会像图15.7那样。同样的，网络中也不会有中心节点和小的集群。

为什么现实世界中的网络会具有这些特征呢？这是网络科学的主要问题，目前基本上已经通过建立网络的发展模型解决了。其中有两类模型被深入地进行了研究，分别是小世界网络（small—world　networks）和无尺度网络（scale—free networks）。

小世界网络

米尔格兰姆的实验也许不能证明我们的社会是一个小世界，但我的社会网络（图15.6）的确是个小世界。从一个节点出发，用不了几步就能到达其他任何节点。Gar只需3步就能到达Charlie，John只需4步就能到达Xiao，虽然他们从未谋面（据我所知是这样）。在我的网络中人们最多相隔4度。

应用数学家和社会学家邓肯·瓦特与应用数学家斯托加茨率先从数学上定义了小世界网络的概念[229]，并且研究了怎样的网络结构会具有这种特性。（他们对网络的抽象研究的灵感来源出人意料：来自对蟋蟀如何同步鸣叫的研究。）瓦特和斯托加茨从一个最简单的“规则”网络开始：由60个节点组成的一个环，如图15.8所示。每个节点与相邻的两个节点相连，像一个初等的元胞自动机。为了确定网络的“小世界”程度，瓦特和斯托加茨计算了网络的平均路径长度。两个节点之间的路径长度就是两个节点之间最短路径的边的数量。平均路径长度则是网络中所有节点对之间的路径长度的平均值。结果图15.8中的规则网络的平均路径长度为15[1]。如果玩传话游戏，要与坐在对面的人沟通会需要很长时间。

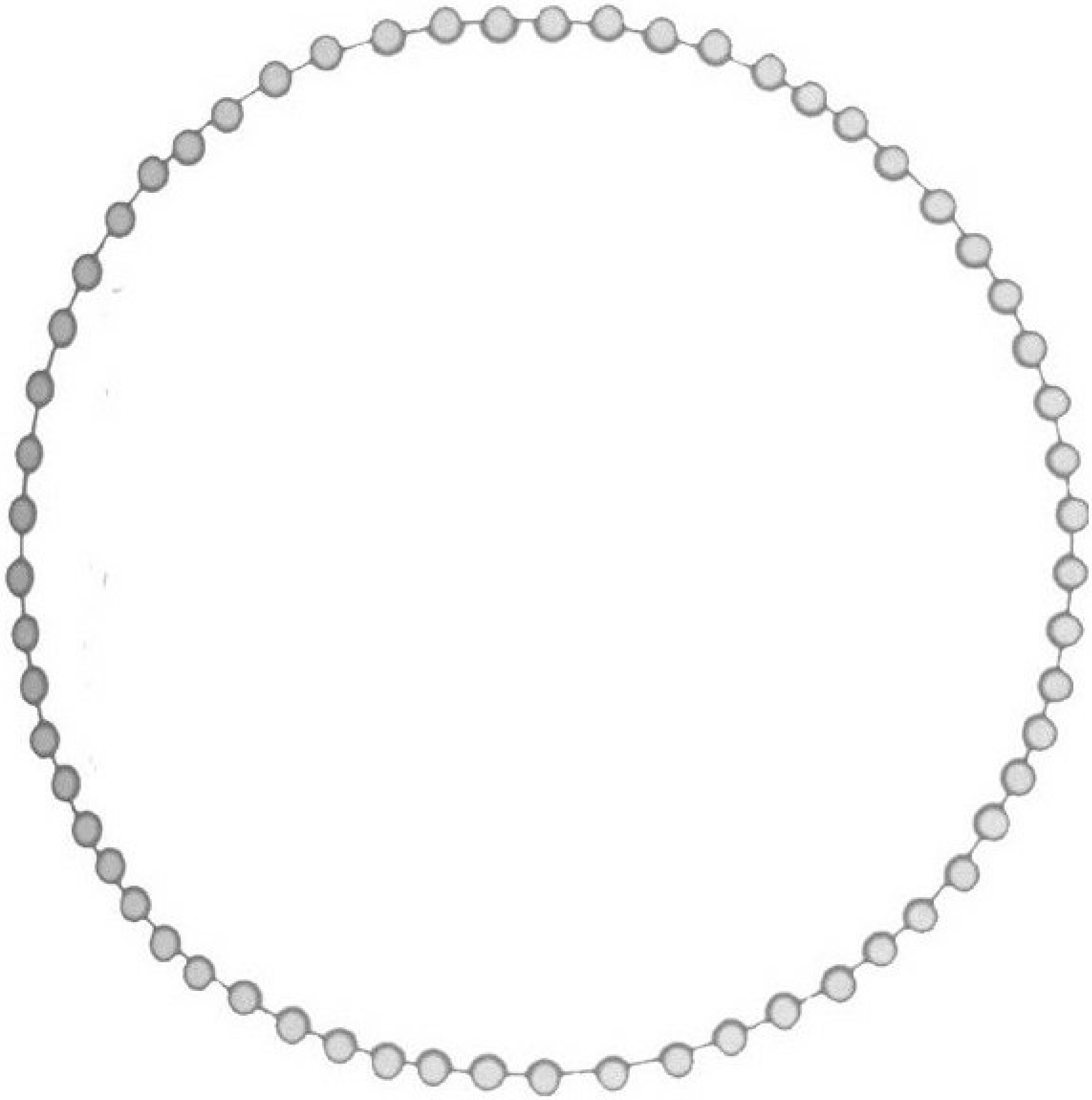


图15.8规则网络的例子。这个网络是由节点组成的一个环，每个节点都与相邻的两个节点相连

瓦特和斯托加茨想知道，如果我们对这样的规则网络稍加改动，将少量与相邻节点连接的边改成长距离连接，平均路径长度会受到怎样的影响呢？他们发现，影响相当剧烈。

图15.9是对图15.8网络中5%的边（3条）进行重连后得到的网络，重连时3条边的一端被解开，重新连接到一个随机选择的节点上。

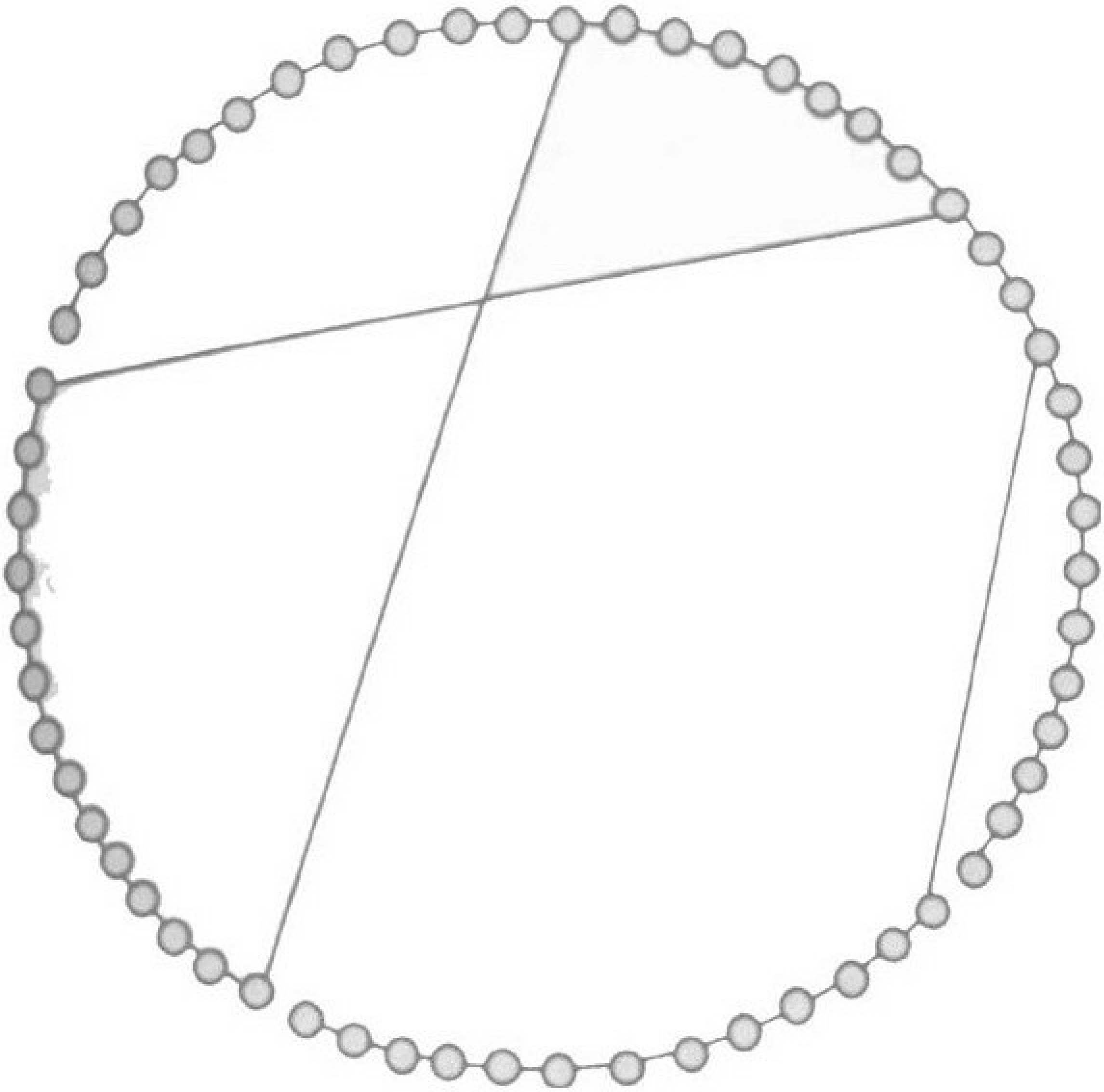


图15.9　将3条边随机重连，使得图15.8的规则网络变成了小世界网络

重连后的网络与原来的规则网络的边数量一样多，但是平均路径长度一下就降到了9左右 [230] 。瓦特和斯托加茨发现，节点数量越多，这个效应越明显。例如，如果是有1000个节点的规则网络，平均路径长度是250，如果5%的边重连，平均路径长度会一下降到20。瓦特说：“只需很少的随机连接就能产生很大的效应 [231] ……不管网络的规模多大，前5个随机重连会将平均路径长度平均减少一半。”

这解释了小世界性 [232] ：一个网络如果只有少量的长程连接，相对于节点数量来说平均路径却很短，则为小世界网络。小世界网络也经常

表现出高度的集群性：任选3个节点A、B、C，如果节点A与节点B和C相连，则B与C也很有可能相连。这在图15.9中不明显，因为这个网络中大部分节点都只与两个相邻节点相连。但如果网络更贴近真实，也就是说节点与更多节点相连，则集群性会很高。我自己的社会网络就是一个例子——我的朋友的朋友也很有可能是我的朋友。

瓦特和斯托加茨还研究了3个真实世界中的网络，结果表明它们都具有小世界性。一个是电影演员网络。在这个网络中，节点代表演员；如果两个演员出现在同一部电影中，例如汤姆·克鲁斯和马克斯·冯·赛多（《少数派报告》，*Minority Report* ），卡梅隆·迪亚兹和朱莉娅·罗伯茨（《我最好朋友的婚礼》，*My Best Friend's Wedding* ），则相应的两个节点就相互连接。这个网络因著名的“凯文·贝肯游戏”[233]（Kevin Bacon game）而受到关注，游戏参与者尝试在网络中寻找任意一位电影演员与多产的电影明星凯文·贝肯的最短路径。一般来说，如果你是演电影的，与凯文·贝肯之间的路径很长，就说明你在演艺界混得不好。

第二个例子是美国西部电网。网络中的节点代表电网的主要组成部分：电厂、变压器、变电站。边代表它们之间的高压输电线。第三个例子是线虫的神经网络，节点代表神经元，边则代表神经元之间的连接。（瓦特和斯托加茨很幸运，神经学家已经绘制出了这种低等生物的所有神经元和连接。[234]）

很难想象电影明星与电力系统之间存在共性，更不要说线虫的脑神经，但瓦特和斯托加茨的研究表明它们实际上都是小世界网络，平均路径很短，具有高度的集群性。

瓦特和斯托加茨1990年发表的著名文章——《“小世界网络”的集体动力学》——引发了网络科学的研究浪潮。科学家们在现实世界中发现了越来越多的小世界网络，在下一章我们将介绍其中一部分。自然、社会和技术演化产生的许多生物、群体和产品似乎都具有这种结构。这是为什么呢？有假说认为至少有两种相互矛盾的选择压力导致了这种结果：在系统内快速传播信息的需要，以及产生和维持可靠的远程连接的高成本。小世界网络具有较短的平均路径长度，同时又只需相对较少的长程连接，从而解决了这两个问题。

进一步的研究表明，根据瓦特和斯托加茨提出的方法——从规则网

络开始，对一小部分边进行随机重连——产生的网络与许多真实世界网络的度分布并不一样[235]。很快不同的网络模型就被提了出来，其中就包括无尺度网络——一种更类似现实世界网络的小世界网络。

无尺度网络

你肯定搜索过万维网，你也很有可能将谷歌作为你的搜索引擎。（如果你读这本书的时间离我写书的2008年已经过去了很久，流行的也许是新的搜索引擎。）在谷歌出现之前，搜索引擎的做法是在一张索引上搜索你查询的单词，索引将所有可能的英文单词对应到包含有这个单词的网页的列表。比如，如果你用“apple（苹果）records（唱片、记录）”这两个单词进行搜索，搜索引擎会列出包含这两个单词的所有网页，根据这些单词的出现次数进行排序。你可能会得到华盛顿州苹果的历史价格网页，或是塔斯马尼亚大苹果赛的最快时间记录，或是甲壳虫乐队1968年的著名唱片的网页。当时在一大堆不相关的网页中寻找你想要的信息是一件让人充满挫折感的事情。

20世纪90年代，谷歌改变了这一切。谷歌提出了一种革命性的思想，用一种称为“网页排名（Page Rank）”的方法对网页搜索结果进行排序。其中的思想是网页的重要性（和可能的相关性）与指向这个网页的链接数量（入连接的数量）有关。例如，在我写下这些的时候，《美国西部果农》报道2008年苹果价格的网页[236]有39个入连接，关于塔斯马尼亚大苹果赛信息的网页[237]有47个入连接。甲壳虫乐队官网（www.beatles.com）有大约27000个入连接。在搜索“apple records”时这个网页位于列表前面。在搜索到的近100万个网页中，另两个网页则远远排在后面。最初的网页排名的思想非常简单，但它极大地改善了搜索引擎，与搜索单词最相关的网页通常都位于列表的前面。

如果我们将万维网看作一个网络，节点是网页，边是网页之间的超链接，我们就能发现网页排名之所以有效是因为这个网络具有特定的结构：同典型的社会网络一样，大部分网页为低连接度（入连接相对较少），极少部分网页具有高连接度（入连接相对较多）。此外，网页之间的入连接数量差别很大，这样才使得排名有意义——能真正对网页进行区分。换句话说，万维网具有前面描述的那种度分布和中心节点结构。而且它也具有高度的集群性——一些网页“群体”内部相互连接。用网络科学的术语说，万维网是无尺度网络。这在最近的复杂系统研究中

是最热门的话题，因此我们来稍加深入地了解一下万维网的度分布，以及无尺度到底指的是什么。

万维网的度分布

怎样分析万维网的度分布呢？万维网连接有两种：入连接和出连接。如果我的网页有一个链接指向你的网页，而你却没有指向我，则我有一个出连接而你则有一个入连接。我们必须明确考虑的是哪种连接。最初的网页排名算法只关注入连接而忽略出连接——在我们的讨论中也是一样。我们将网页的入连接数量称为网页的入度（in—degree）。

这样问题就是万维网的入度分布是怎样的？要考虑所有的网页和入连接极为困难——并不存在完整的列表，而且不断有新的连接产生，旧的连接消失。不过一些网络学家还是通过采样和巧妙的网络爬虫（Web—crawling）技术得到了近似结果。对网页总数的估计各种各样；2008年时，我看到的估计从1亿到100亿都有，而且显然网页数量还在迅速增长。

几个研究团体都发现网页入度分布可以用非常简单的规则来描述：具有某一入度的网页数量大致正比于入度的平方的倒数。如果用字母k表示入度。则（在文献中对于k的指数的具体数值有些分歧，但是都接近2——细节见注释）这个规则实际上只适用于[238]入度（k）为数千或更大的情形。

入度为**k**的网页数量正比于$\mathbf{1/k^2}$。

为了解释为何万维网是“无尺度”，我用三种不同的尺度画出了遵循上面的规则的入度分布（图15.10）。第一幅图画了9000个入度对应的分布，从1000开始，这是规则开始变得相当精确的地方。类似于图15.7，横轴为1000—10000的入度值，纵轴标示的长条高度则是各入度值的频率（具有相应入度的网页数量）。长条太多以至于形成了一整块黑色区域。

图中给出的并不是频率的真实值，因为我想强调图形的形状（据我所知，也没有人对实际频率有很好的估计）。图中可以看到入度k为1000的网页相对较多，随着入度增大，频率下降很快。在k=5000和

k=10000之间，网页数量已经很少了，对应的长条高度已基本接近0。

如果我们改变尺度，将这片“基本接近0”的区域放大，会怎么样呢？第二幅图绘制了k=10000到k=100000之间的入度分布。我改变了图形的尺度，让k=10000处与前一幅图中k=1000处的长条高度相等。在这个尺度下，入度为k=10000的网页数量相对较多，在k=50000和k=100000之间的长条高度则“基本接近0”。

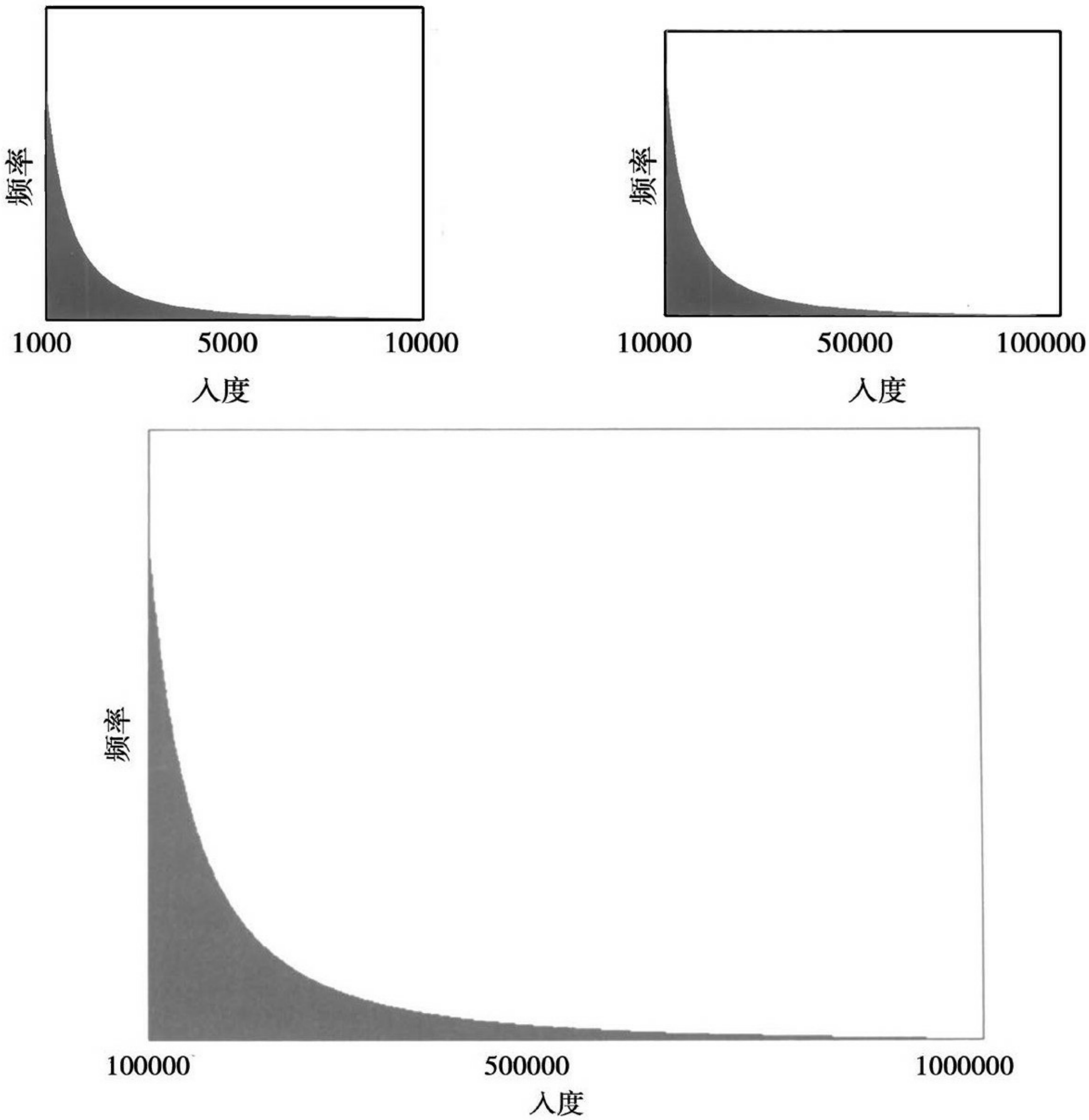


图15.10　三种不同尺度下网页入度分布的近似形状

但还有一点很惊人——除了坐标轴上的数字不同，第二幅图与第一幅图是一样的。第一幅图画了9000个值的分布，而第二幅图画了90000个值的分布，足足大了一个数量级。

第三幅图中可以看到在更大的尺度上仍然有相同的现象。k为100000到1000000之间，画出的分布形状仍然是一样的。

这样的分布被称为是自相似的，因为不管在哪种尺度上进行绘制，形状都是一样的。说得更专业一点，就是“在不同尺度下具有不变性”。这就是无尺度一词的由来。自相似这个词可能会让你觉得似曾相识。因为在第7章讨论分形时我们见过它。这里与分形确实有一些联系，到第17章我们再来详细讨论。

无尺度分布和钟形曲线

无尺度网络没有“特征尺度”。要解释这一点，最好的办法是将无尺度分布与另一种著名的分布——钟形曲线——进行比较。

假设我们绘制全世界成年人身高的分布。世界上最矮的（成年）人大约是70厘米，最高的人则大约是270厘米。成年人的平均身高大约是165厘米，而且大部分人的身高介于150—200厘米之间。人类身高分布大致像图15.11那样。画出来的曲线像一座钟，钟形曲线也因此而得名。很多东西的分布都接近钟形曲线——身高、体重、考试成绩、篮球赛得分、各种物种的数量，等等。自然界中有许多量都遵循这种分布，因此钟形曲线也被称为正态分布。正态分布有特定的尺度——比如，身高是70—270厘米，考试成绩是0—100分。在正态分布中，平均值同时也是频率最高的值，例如165厘米既是身高平均值也是最常见的值。大部分取值与平均值相差不大——分布相当单一。如果网页的入度值是正态分布，网页排名就不会起作用，因为几乎所有网页的入连接都差不多。甲壳虫乐队的官网与其他包含“apple records”的网页的入连接数量就会相差不大，因而无法用来区分可能的相关程度。

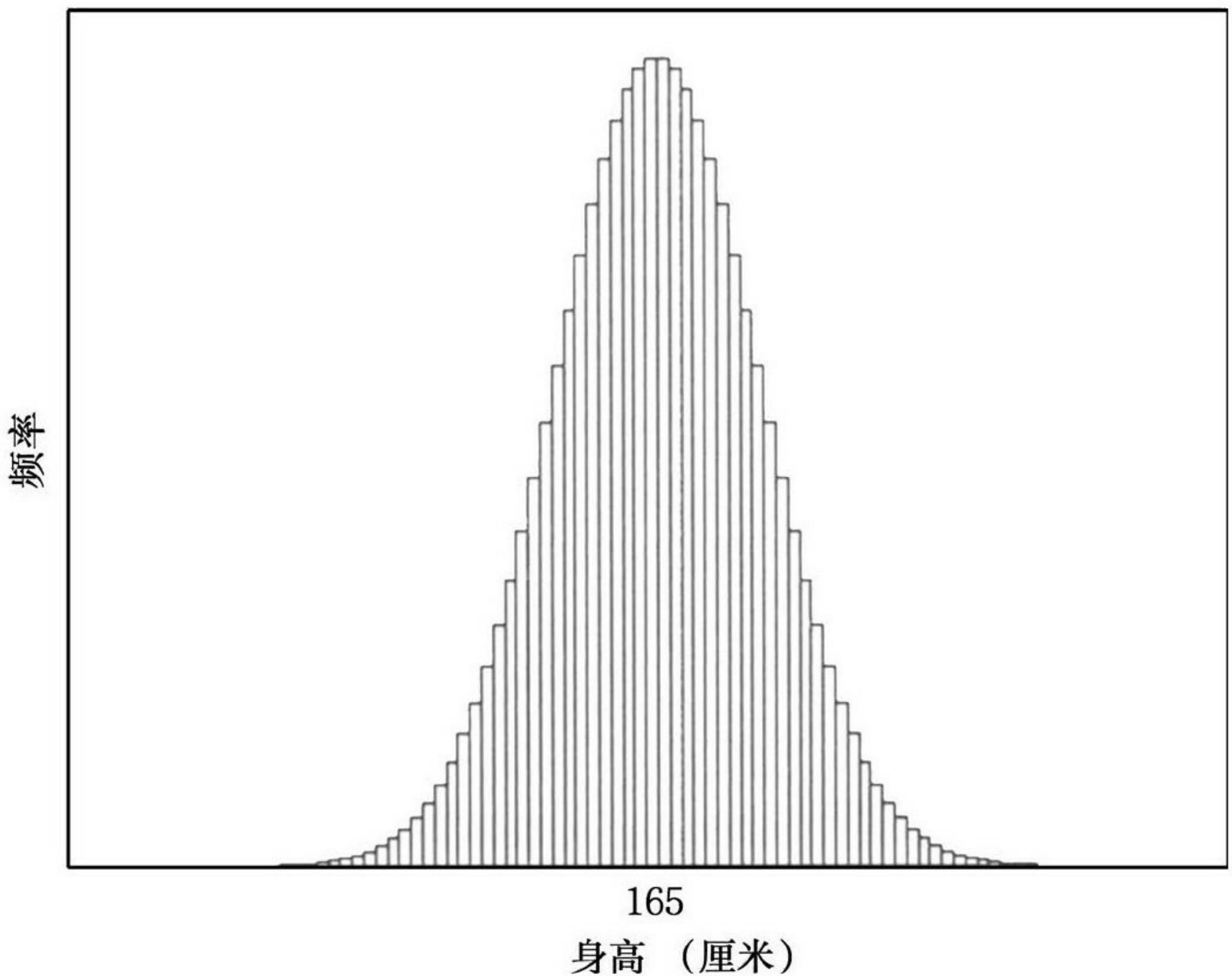


图15.11 人类身高的钟形曲线分布

幸运的是（对谷歌的股东来说尤其如此），网页的度分布是无尺度而不是钟形曲线。无尺度网络有4个显著特征：①相对较少的节点具有很高的度（中心节点）；②节点连接度的取值范围很大（度的取值多样）；③自相似性；④小世界结构。所有的无尺度网络同时也具有小世界特性，但不是所有具有小世界特性的网络都是无尺度网络。

说得专业一点，无尺度网络一定遵循连接度幂律分布。前面已经讲了，网页的入度分布大致是：

入度为**k**的网页数量正比于**$1/k^2$**。

你可能还记得高中数学学过$1/k^2$ 也可以写成k^{-2} 。即“指数为-2的幂律分布”。类似的，$1/k$即（k^{-1} ）就是“指数为-1的幂律分布”。幂律分布

的形式一般为x^d，其中x是入度这一类的量。描述这种分布的关键是指数d；不同的指数会产生不同的分布。

在第17章我们再深入讨论幂律分布。现在只需记住无尺度网络遵循连接度幂律分布。

网络稳健性

无尺度网络有一个非常重要的特性，在节点被删除时具有稳健性。也就是说，如果随机删除一些节点，不会改变网络的基本特性：仍然会有多样的度分布、很短的平均路径以及很高的集群性，即使删除的节点很多也不会有什么变化。原因很简单：如果随机删除节点，则极有可能删除的是低连接度的节点，因为网络中绝大部分节点都是低连接度节点。删除这种节点对总体的度分布和路径长度的影响很小。万维网就是这样，网络上不断有计算机出故障或是被移除，但是这对万维网的运转不会有明显影响，也不会改变其平均路径长度。类似的，网页和链接也在不断被删除，但网上冲浪不会受到什么影响。

不过，这种稳健性是有代价的：如果删除了中心节点，网络就有可能会失去无尺度特性，并且无法正常运转。例如，芝加哥（航班网络的中心节点）的暴风雪可能会导致全国大面积的航班延误或取消。谷歌出故障会对整个万维网形成很大冲击。

总而言之，无尺度网络对节点的随机删除具有稳健性，但如果中心节点失效或是受到攻击就会非常脆弱。

下一章我们将讨论几个具有小世界和无尺度特性的真实网络，以及一些对它们的形成进行解释的理论。

第16章　真实世界中的网络

网络的思想显然深入人心。我在谷歌学术上搜索了一下，从2003年到我写书时的5年里，关于小世界或无尺度网络的论文超过了14000篇，仅去年一年就有近3000篇。我浏览了一下前面约100篇论文的标题，发现涉及11个不同的学科，既包括物理和计算机，也包括地质学和神经科学。我相信如果继续往后看的话，涉及的学科还会更多。

这一章我们来看几个真实世界中的网络，然后讨论一下网络科学的进展对各学科的学者思维方式的影响。

真实世界中的网络

大脑

一些研究团队发现的证据表明大脑具有小世界特征。大脑在几个不同的描述层面上都可以视为网络。例如，将神经元作为节点，突触作为边，或者将整个功能区作为节点，将功能区之间的大尺度连接（神经元连接群）作为边。

前一章曾提到，神经学家已经完整绘制了线虫的脑神经网络，并发现线虫的脑是小世界网络。最近，神经学家又绘制出了猫、恒河猴等动物甚至人类的一些高级功能脑区的连接结构，[239] 并且发现这些网络同样具有小世界特性。

为什么进化喜欢具有小世界特性的大脑网络呢？弹性可能是一个重要原因：我们知道神经元会不断死去，但幸运的是，大脑仍然能正常运转。大脑的中心节点则是另一回事：比如海马区（负责短时记忆的网络的中心），如果受到击打或是疾病侵袭，后果将会是毁灭性的。

另外，研究者还猜测，连接度的无尺度分布使得大脑可以在两种大脑行为之间达成最佳妥协：信息在视觉皮层或语言区等局部区域的处理，以及信息的全局处理，例如视觉皮层的信息传递到语言区，或者反过来。

如果每个神经元或是所有功能区相互之间都有连接，在这些连接上传递信号耗费的能量将大得惊人。进化可能选择了能效更高的结构。另外，如果那样大脑的体积也要大得多。另一方面，如果大脑中没有长程连接，则不同区域之间的通信会困难得多。人类的脑容量——以及相应的颅骨大小——似乎在大小上形成了精妙的平衡，要大到足以进行复杂的认知，同时又要小到可以让母亲生下来。有观点认为正是小世界特性让这种平衡得以达成。

科学家们普遍认为，同步——神经元群不断同时激发——是大脑中信息高效传播的主要机制，而小世界结构极大地促进了这种同步的产

生。

基因调控网络[240]

第7章曾讲过，人类大约有25000个基因，与拟南芥的基因数量差不多。人类之所以比植物复杂，不在于基因数量，而在于基因如何相互作用。

有很多基因的作用就是调控其他基因——即决定受调控的基因是不是表达。一个著名的基因调控的例子就是大肠杆菌乳糖代谢的控制。这种细菌通常以葡萄糖为食，但也能代谢乳糖。细胞要代谢乳糖必须有三种特定的蛋白酶，每种都由单独的基因编码。我们先称这些基因为A、B和C。另外还有一种乳糖抑制蛋白（lactose repressor），能够与基因A、B、C结合，从而关闭这些基因。如果在细菌的周围没有乳糖，乳糖抑制蛋白就会不断产生，从而不发生乳糖代谢。但如果细菌突然发现周围没有葡萄糖，而乳糖又很多，乳糖分子就会与乳糖抑制蛋白结合，让其离开基因A、B和C，这些基因就能产生酶，从而进行乳糖代谢。

有些调控机制还要复杂精巧得多，这些调控作用是遗传复杂性的精髓。早在20世纪60年代，考夫曼（Stuart Kauffman）在研究调控机制时就利用了网络的思想（详见第18章）。最近，网络科学家和遗传学家又一起合作发现，至少有一些调控网络接近于无尺度。在调控网络中，节点代表单独的基因，边则代表基因之间的调控关系（如果有的话）。

稳健性对于基因调控网络也很重要。基因转录和调控的过程远不是完美的；它们难免犯错，而且经常受病毒等病原体的侵袭。无尺度结构能让系统基本上不受这些错误影响。

代谢网络

第12章曾讲过，大部分生物的细胞都有上百种代谢途径，代谢途径之间相互作用，形成代谢反应网络。巴拉巴西和他的同事仔细研究了43种生物的代谢网络结构，[241] 发现它们都符合幂律分布——也就是说，是无尺度网络。代谢网络中的节点是化学反应物——化学反应的原料和产物。如果某种反应物参与了生成另一反应物的反应，就认为前者连接到后者。例如，在糖酵解这种代谢途径的第二步，葡萄糖—6—磷酸

（glucose—6—phosphate）生成果糖—6—磷酸（fructose—6—phosphate），因此在网络中前一反应物有边连接到后一反应物。

既然代谢网络是无尺度的，就有少数中心节点是许多种反应的产物，涉及许多不同的反应物。结果发现，在所研究的所有生物中，这些中心节点代表的化学物质基本都是一样的——对生命最重要的化学物质。有假说认为，代谢网络之所以演化出无尺度特性，是为了确保代谢的稳定性，并优化不同反应物之间的“通信”。

流行病

20世纪80年代初，艾滋病流行的早期阶段，美国疾控中心（Centers for Disease Control）的流行病学家发现，有一个群体将艾滋病毒传播到了全世界许多城市的男同性恋，这群人中包括一位加拿大空乘——盖坦·杜格斯（Gaetan Dugas）。杜格斯后来被媒体蔑称为“零号病人”，北美第一个艾滋感染者，认为他对艾滋病毒在美国等地的传播负有责任。虽然后来的研究否认了杜格斯是北美的传染源，[242] 但毫无疑问杜格斯感染了许多人，他承认自己每年都有上百个不同的性伴侣。用网络的术语说，杜格斯是性关系网络的中心节点。

研究性传播疾病的流行病学家经常需要研究性关系网络，这个网络中的节点代表人，边则代表人之间的性伴侣关系。最近，一个由社会学家和物理学家组成的团队分析了瑞典性行为调查数据，[243] 结果发现得出的网络为无尺度结构；其他对性关系网络的研究也得出了类似结论[244]。

在这种情形下，移除中心节点就对我们有利。专家建议，安全性行为宣传、疫苗接种等干预措施应当主要针对这类中心节点。

但是得不到性关系的数据，绘制不出整个网络，又如何能识别出中心节点呢？

另一个网络科学家团体提出了一个巧妙而简单的方法[245]：从风险人群中随机选取一组人，让他们每人提供一位性伴侣的名字。然后给这些性伴侣接种疫苗。性伴侣很多的人出现在名单中的概率会很高，从而通过这种方案被接种疫苗。

当然这种方法也可以用到其他场合，用来进行“中心节点打靶”，比如对付通过电子邮件传播的病毒：对于这种情况，杀毒应当重点针对邮件通信录很长的用户的计算机，[246]而不是寄希望于所有计算机用户都能查杀病毒。

生态与食物网

在生态学中，食物链的传统概念已经转变成食物网（food web）的概念，食物网中的节点代表物种或物种群；如果物种B是物种A的食物，就有一条边从节点A连接到节点B。图16.1展示了一个食物网的简单例子。

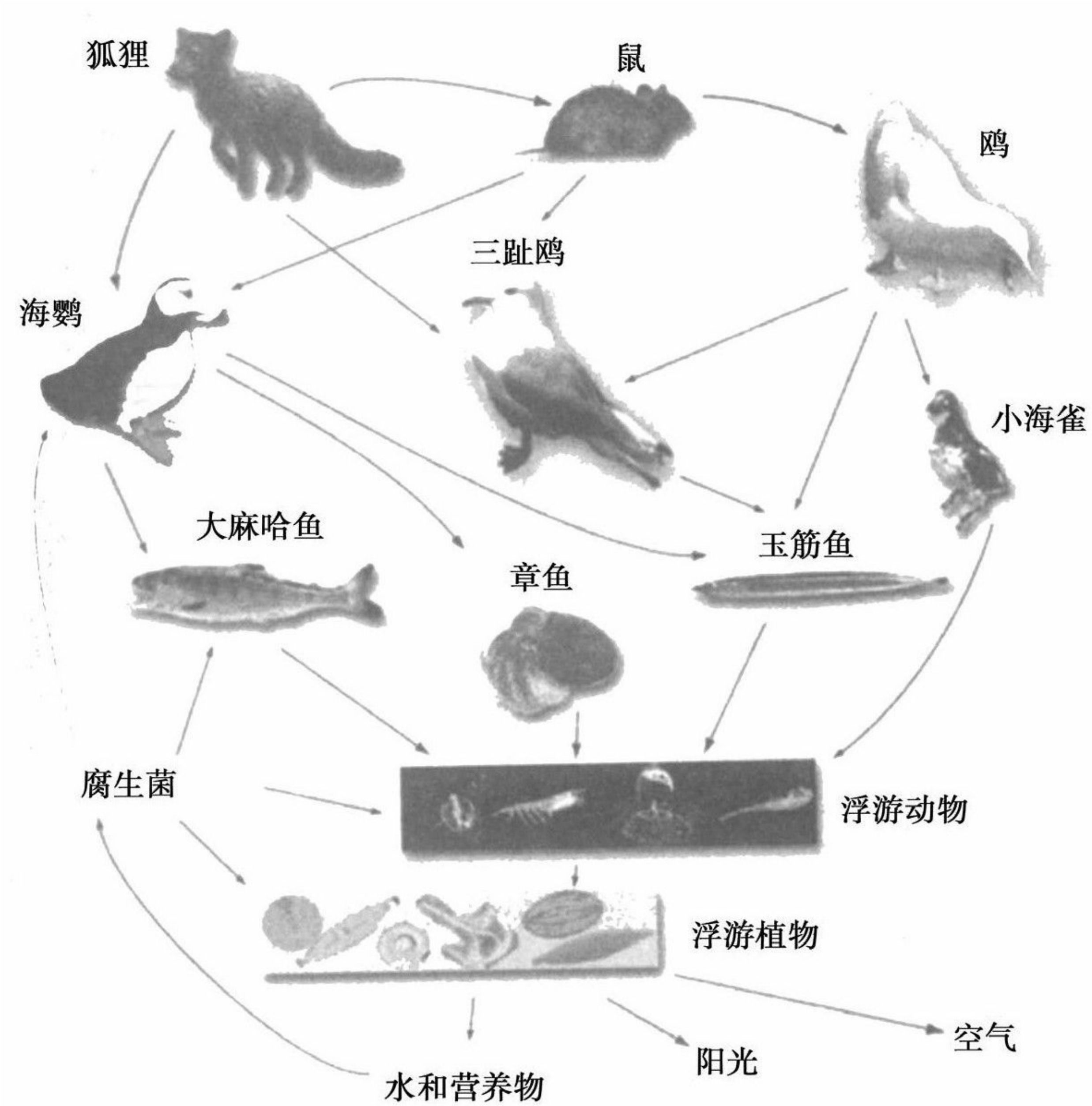


图16.1　食物网示例[图例来自美国地质勘探局阿拉斯加科学中心（USGS Alaska Science Center, http：//www.absc.usgs.gov/research/seabird_foragefish/marinehabitat/home.html）]

绘制各种生态系统的食物网一直是生态学的重要内容。最近，科学家们开始用网络科学来研究食物网，深入理解生物多样性以及破坏生物多样性会带来的可能后果。

一些生态学家认为（至少部分）食物网具有小世界特性，其中一些还具有连接度无尺度分布，这种特点可能使食物网在面对物种的随机灭绝时具有一定的稳健性。另一些生态学家则不同意食物网具有无尺度结

构，生物学界最近对这个问题有很多争议，[247] 主要是对如何解读真实数据难以达成共识。

网络思想的意义

网络思想对许多科技领域都有影响，前面只是一小部分例子。连接度无尺度分布、集群性和存在中心节点是共同的主题；这些特点使得网络具有小世界的通信能力，并且在随机删除节点时具有稳健性。所有这些特点都有助于理解科学和技术领域的复杂系统。

在科学领域，网络思想为描述自然界复杂系统的共性提供了新的语言，也使得从不同领域得到的知识能相互启发。就其本身来说，网络科学正是它自己所说的那种中心节点——它使得本来相隔遥远的学科变得很近。

在技术领域，网络思想为许多困难问题提供了新的思路，例如，如何让网络上的搜索变得高效，如何控制流行病，如何管理大型组织，如何保护生态系统，如何应对威胁身体中的复杂系统的疾病，如何应对现代犯罪和恐怖组织，以及在更高层面上，自然、社会和技术网络有怎样内在的稳健性和脆弱性，又应当如何利用和保护这种系统。

无尺度网络是如何产生的[248]

没有谁有意识地将万维网设计成无尺度分布。万维网的连接度分布，同前面提到的所有网络一样，是网络在形成过程中涌现的产物，是由网络的生长方式决定的。

1999年，物理学家巴拉巴西和艾伯特提出了一种网络生长机制[249]——偏好附连（preferential attachment），用来解释大部分真实世界网络的无尺度特性。其中的思想是，网络在增长时，连接度高的节点比连接度低的节点更有可能得到新连接。直观上很明显。朋友越多，就越有可能认识新朋友。网页的入度越高，就越容易被找到，因此也更有可能得到新的入连接。换句话说就是富者越富。巴拉巴西和艾伯特发现，偏好附连的增长方式会导致连接度无尺度分布。（他们当时不知道，这种机制以及所产生的幂律在以前至少被独立发现过三次[250]。）

科技文献引用网络的增长[251]是偏好附连效应的一个例子。在这个网络中节点是科技文献；一篇论文如果被另一篇论文引用，就得到一条入连接。因此论文被引用的次数越多，连接度就越高。人们一般认为被引用次数越多，论文就越重要；在科学界，这个指标会决定你的职位、加薪等等。不过，偏好附连似乎经常在其中扮演重要角色。设想你和科学家乔各自独立地就同一个问题写了很出色的论文。如果我在我的论文中碰巧引用了你的文章，却没有引用乔的，其他人如果只读了我的文章就很有可能会引用你的文章（经常是读都没读）。其他人如果读到了他们的文章，也会更有可能引用你的而不是乔的文章。局势会越来越有利于你，不利于乔，尽管乔的论文和你的论文质量一样好。偏好附连机制会导致作家格拉德威尔（Malcolm Gladwell）所说的引爆点[252]（tipping points）——论文引用、时尚流行等过程通过正反馈循环开始剧烈增长的点。另外，引爆点也可以指系统中的某处失效引发系统全面加速溃败，后面我们将讨论这种情况。

幂律以及对其的质疑

前面我们说了，无尺度网络的连接度幂律分布能使系统稳健、通信迅速，这也使得这种网络在自然界很普遍，而它们的形成机制主要是偏好附连。这个观念给科学家研究其他问题带来了新的思路。

不管一种思想如何具有吸引力，科学家们对之都天生抱有怀疑， 尤其是新提出的思想，还没有被怎么检验过，对于那种声称对很多学科都具有普适性的思想更是如此。这种怀疑态度是健康的，它是科学能够进步的关键。因此，并不是所有人都加入了网络科学阵营，事实上很多人认为，对于网络科学在复杂系统研究中的意义，一些观点过于乐观。下面我们来看看这些怀疑观点。

1.幂律和无尺度分布被滥用。对于真实世界中的网络，一般都很难得到好的数据。例如，巴拉巴西和他的同事在研究代谢网络时用的是网上的数据库，其中的数据由世界各地的生物学家提供。这类生物数据库对于研究很有帮助，但同时也必然是不完善的，还有很多错误。巴拉巴西和他的同事只能靠统计和曲线拟合来计算各种代谢网络的度分布，这种方法还存在问题，但是分析真实数据时大部分用的都是这种方法。用这种方法确定为“无尺度”的一些网络后来又被发现其实不是无尺度分布。[253]

就像哲学家和生物史学家凯勒（Evelyn Fox Keller）说的："现在幂律的普遍性可能被高估了 [254] 。"物理学家和网络学家沙利兹（Cosma Shalizi）的话说得更不客气："我们对幂律的迷恋是一种耻辱 [255] 。"就在我写下这些的时候，对于真实世界的网络是不是确实是无尺度分布仍然存在很多争议。

2.即使网络确实是无尺度分布，也有很多可能会导致网络的连接度幂律分布；并不一定非得是偏好附连。沙利兹一针见血地指出："产生幂律的方法很多，每种都有道理 [256] 。"我在圣塔菲研究所的时候，几乎每隔一天都有讲座介绍产生幂律分布的新机制 [257] 。有一些与偏好附连类似，有一些则相差很大。很难说到底是哪种机制导致了在真实世界中观察到的幂律。

3.网络科学的模型过度简化，假设前提不现实。小世界和无尺度网络模型也仅仅是模型而已，模型中的简化假设在现实世界中的网络也许不成立。建立这些简化模型的目的是希望能抓住所研究的现象的某些方面。前面我们看到，这两个网络模型，尤其是无尺度模型，确实抓住了一些东西，大量真实系统中的度分布、集群以及稳健性（虽然前面第1点认为适用面可能不像想象的那样广）。

不管怎样，网络的简化模型就其本身来说，还是无法解释真实世界中的网络的一切。不管是小世界还是无尺度模型，所有节点都被当作是一样的，除了连接度；所有的边的类型和强度也是一样的。在真实世界中的网络不是这么回事。例如，在我的社会网络中（图15.6有简化模型），一些代表友情的边就比其他边的强度要大些。Kim和Gar都是我的朋友，但是我和Kim关系更好，因此我也更有可能把我个人的重要事情告诉她。而且，Kim是女人，而Gar是男人，这也使得我更倾向于信赖Kim。同样的，比起Kim来，我的朋友Greg要懂数学一些，因此如果我想讨论纯数学方面的问题，我就更有可能找Greg而不是Kim。边和节点类型的区别，以及边的强度，对于信息在网络上的传播有很大影响，而简化的网络模型无法抓住这种影响。

网络中的信息传播和连锁失效

事实上，理解信息在网络中的传播方式是网络科学现在面临的最重要的问题。到目前为止我们讨论的都只是网络的结构——例如，静态的

度分布——还没有讨论网络中信息传播的动态行为。

“网络中的信息传播”指的是什么呢？这里的信息一词指的是节点之间的所有交流。信息的传播包括例如谣言、流言、流行时尚、思想、流行病（通过病毒传播）、电流、万维网上的数据包、神经递质、卡路里（在食物网中传递），以及一种更普遍的网络传播现象——“连锁失效（cascading failure）”。

连锁失效现象的存在促使人们关注网络中的信息传播以及其如何受网络结构影响。网络中的连锁失效是这样一个过程：假设网络中每个节点都负责执行某项工作（例如传输电力）。如果某个节点失效了，它的工作就会转移到其他节点。这有可能会让其他节点负荷过重从而失效，又将它们的工作传递到其他还未失效的节点，这样不断发展。结果是失效如同加速的多米诺骨牌一样扩散，从而让整个网络崩溃。

连锁失效的例子在现实网络世界中很常见。下面是新闻里最近报道的两个例子：

◆2003年8月：美国中西部和东北部发生大规模断电，是由俄亥俄州一家发电厂发生故障引发的连锁失效导致的。据报道，由于天气过于炎热，[258] 导致电线负荷过高，引起线路下垂，碰到了树枝，触发了线路自动断路，负载被转移到电网其他部分，使得其他部分也因过载而失效。过载失效迅速传播，最后导致加拿大和美国东部5000万居民断电，有些地区断电长达3天。

◆2007年8月：美国海关计算机系统崩溃了近10个小时，[259] 导致17000多名旅客滞留在洛杉矶国际机场。事故是由一台计算机的网卡故障引起的。这个故障很快导致其他网卡也连锁失效，不到1个小时，整个系统都崩溃了。海关职员无法处理到达的国际旅客，其中一些人不得不在飞机上等了5个多小时。

第3个例子表明不仅电力网络会发生连锁失效，公司网络也一样。

◆1998年8—9月：私人金融对冲基金美国长期资本管理公司[260]（Long-Term Capital Management, LTCM）得到数家大型金融公司担保从事风险投资，结果将公司的权益价值几乎赔光。美联储担心它的亏损会导致全球金融市场崩溃，因为为了偿债，LTCM会不得不卖掉大部分

资产，导致股票等有价证券的价格下跌，从而迫使其他公司也抛售资产，导致价格进一步下跌，直至崩溃。1998年9月末，为了防止出现这种局面，美联储召集了其主要债权银行对LTCM进行援助。

前面我们说到，在节点随机失效时，对网络的平均最短路径长度不会有很大影响。这种特性在连锁失效的情况下并不成立，因为一个节点的失效会导致其他节点也失效。连锁失效是“引爆点”的又一个例子，小事件触发加速正反馈，结果小问题导致严重后果。许多人担心黑客和电脑恐怖分子威胁全球网络基础，但连锁失效带来的威胁可能更大。随着我们的社会越来越依赖计算机网络、网络投票机、导弹防御系统、电子银行等等，连锁失效的情况也越来越常见，威胁也越来越大。正如研究这种系统的专家安东诺普洛斯（Andreas Antonopoulos）指出的，“威胁来自复杂性本身”[261]。

因此，对连锁失效及其应对策略的总体研究现在是网络科学最活跃的研究领域。两个影响最大的理论分别是自组织临界性[262]（Self-Organized Criticality, SOC）和高容错性[263]（Highly Optimized Tolerance, HOT）。SOC和HOT理论也提出了不同于偏好附连的机制解释无尺度网络的产生。这两个理论各自提出了一组进化和工程系统连锁失效的普适机制。

上一章介绍的小世界网络和无尺度网络简化模型给许多学科引入了网络的思想，并且建立了独立的网络科学领域。下一步我们将了解网络中信息等因素的动力学。要研究免疫系统、蚁群、细胞代谢（参见第12章）这类网络中的信息动力学，网络科学必须刻画节点和边随着时间和空间不断变化的网络。这将是一个重大挑战，就像邓肯·瓦特说的：“对于网络的动力学之谜，[264]不管是疾病传播、电网崩溃，还是革命的爆发，我们现在面临的网络问题就像海边的鹅卵石一样多。”

第17章　比例之谜

前两章我们看到，网络的思想对许多科学领域都有深刻影响，尤其是生物学。就在不久前，网络思想又为生物学中一个最让人费解的难题提出了新的解答：生物的大小变化时其他属性会如何变化。

生物学中的比例缩放

比例描述的是一个属性改变时，其他相关的属性会如何改变。生物学的比例之谜关注的则是生物在休息时消耗的平均能量——基础代谢率——随着生物体重的变化会如何变化。细胞将食物、空气和光照转化为能量的代谢过程是所有生物系统的关键，因此体重与代谢的关系对于理解生命的运作极为重要。

人们很早就发现，相对于体重大小来说，较小动物的代谢率比较大的动物更快。1883年，德国生理学家鲁伯纳（Max Rubner）尝试从热力学和几何的角度来确定准确的比例关系。第3章曾提到过，代谢过程要将能量从一种形式转化成另一种形式，因此总是会发热。生物的代谢率可以定义为细胞将营养转化为能量的速率，能量用于细胞的运作和生成新细胞。在这个过程中生物会以同样的速率散发热量。因此通过测量生物产生的热量就能推导出代谢率。

如果你之前不知道较小动物的代谢率相对比较大的动物更快，你很可能会想代谢率是不是与体重呈线性比例关系——例如，仓鼠的体重是老鼠的8倍，那么代谢率也应该是老鼠的8倍，或者举个极端点的例子，河马的体重是老鼠的125000倍，那么代谢率也应该是老鼠的125000倍。

问题是，如果这样，仓鼠产生的热量也应该是老鼠的8倍。但是散热要通过表皮，而仓鼠的表皮面积只是老鼠的4倍。这是因为动物的表皮面积并不正比于动物的体重（同样也不正比于体积）。

图17.1说明了这一点，图中老鼠、仓鼠和河马用球体表示。你也许还记得中学几何学过球体的体积公式是（4/3）π乘以半径的立方，其中π≈3.14159。而球体的表面积公式是4π乘以半径的平方。因此“体积与半径的立方呈比例”而“表面积与半径的平方呈比例”。这里“呈比例”的意思也就是“正比于”——也就是说忽略常数（4/3）π和4π。从图17.1中可以看到，仓鼠球体的半径大约是老鼠球体的2倍，因此表面积是老鼠球体的4倍，而体积则是老鼠球体的8倍。河马球体的半径是老鼠球体的50倍（没有按比例画），因此河马球体的表面积是老鼠球体的2500倍，而体积则是老鼠球体的125000倍。可以看到，随着半径增大，表面积的增长要比体积慢得多。因为表面积与半径平方呈比例，而体积与半径立方呈比例，因此我们说“表面积与体积的2/3次幂呈比例”。[265]（推导过程

见注释。）

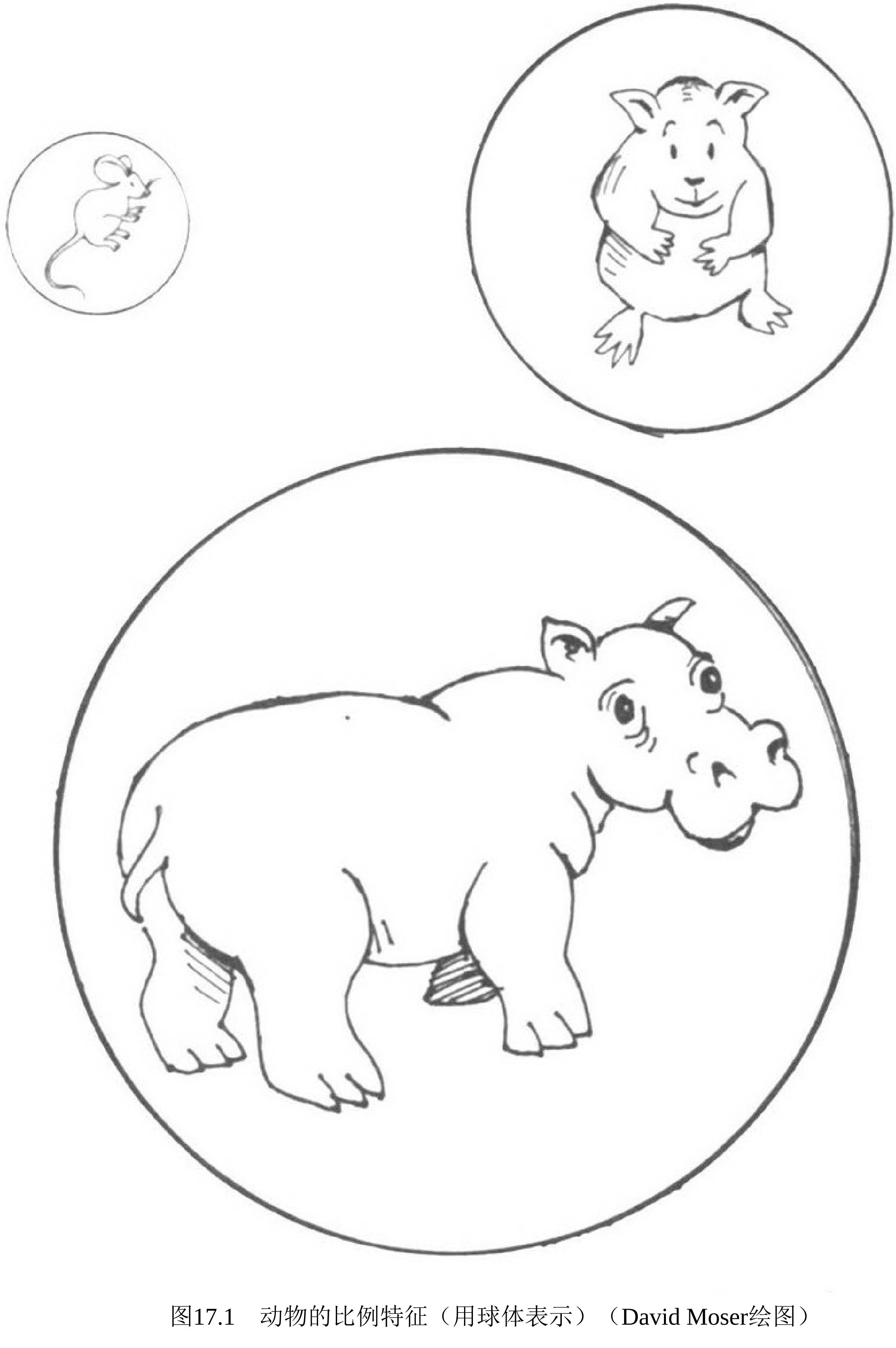


图17.1　动物的比例特征（用球体表示）（David Moser绘图）

体积的2/3次幂的意思就是“体积平方，然后开3次方”。

表皮面积只大4倍，通过表皮发的热却要大8倍，这只仓鼠肯定不是一般的热。同样，河马的表皮面积比老鼠大2500倍，发的热却是老鼠的125000倍。天哪！这只河马恐怕会烧起来。

大自然非常仁慈，它没有这样做。谢天谢地，我们的代谢率与我们的体重不是呈线性比例。鲁伯纳推测，为了安全地发散热量，大自然让我们的代谢率与体重的比例关系同表皮面积一样。他提出代谢率同体重的2/3次幂呈比例。这就是所谓的“表皮猜想（surface hypothesis）”，此后50年这个猜想被广泛接受。唯一的问题是实际数据并不与之相符。

20世纪30年代，瑞典动物学家克莱伯（Max Kleiber）仔细测量了一系列动物的代谢率。他的数据表明代谢率与体重的3/4次幂呈比例。也就是说，代谢率正比于体重$^{3/4}$。你肯定注意到了这就是一个指数为3/4的幂律。这个结果出人意料。指数为3/4而不是2/3，这意味着动物——尤其是较大的动物——的代谢率比人们预想的要高，这也意味着动物比先前根据几何简单预计的要更高效。

图17.2展示了各种动物的这种比例关系。横轴表示体重（单位：千克），纵轴则表示平均基础代谢率（单位：瓦）。图中黑点表示各种动物的实际测量值，直线则表示与体重的3/4次幂呈比例的代谢率曲线。数据与曲线没有精确匹配，但也符合得很好。图17.2是一种特殊的图——专业上称为双对数（或对数—对数）图——图中两条轴都是以10次幂增长。在双对数图上幂律曲线表现为直线，而直线的斜率则等于幂律的指数。[266]（见注释中对此的解释。）

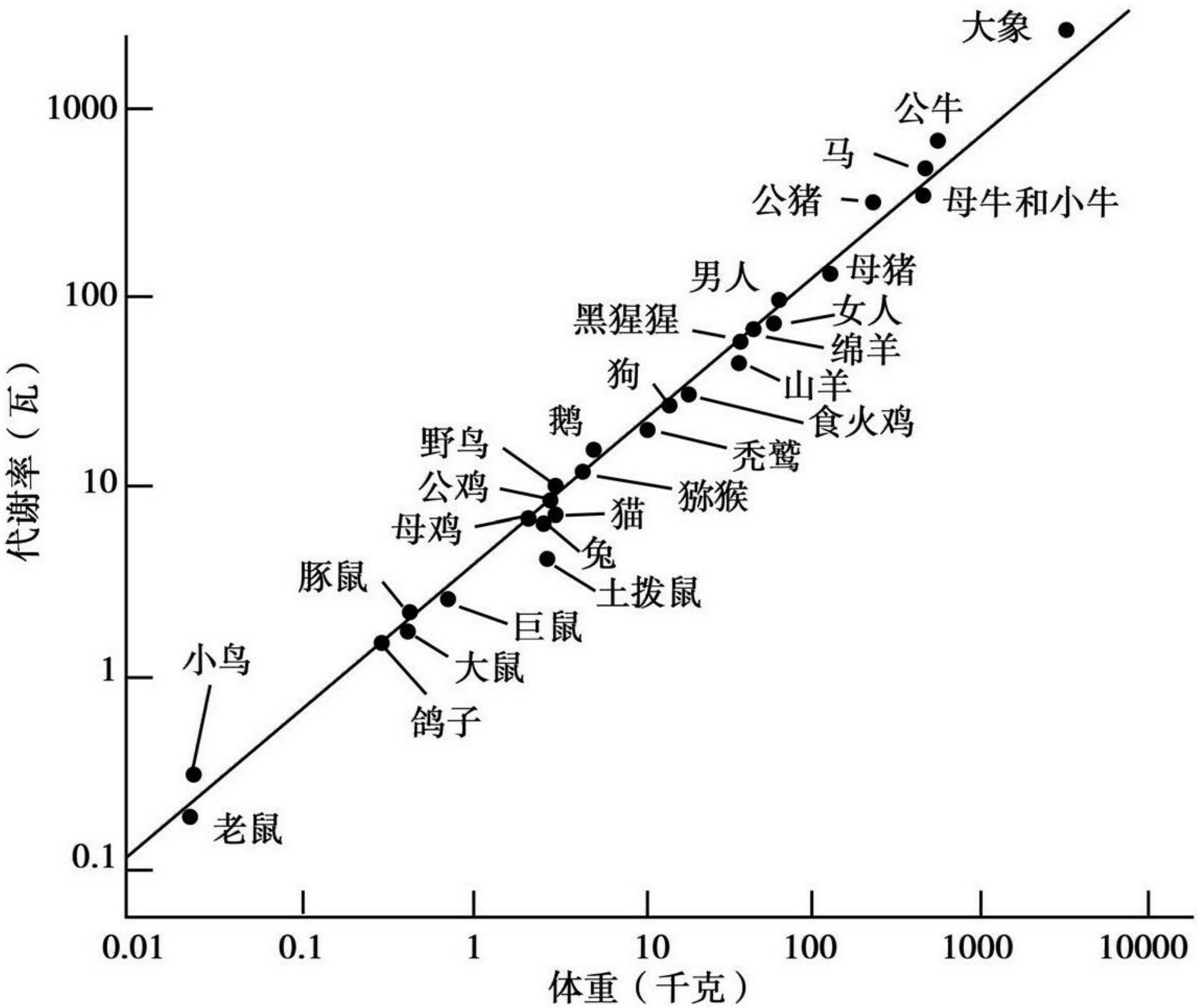


图17.2　各种动物的代谢率与体重的关系[引自施密特—尼尔森的《比例：为什么动物的大小如此重要？》（K.Schmidt-*Nielsen, Scaling*：*Why Is Animal Size So Important*？），剑桥大学出版社1984年出版。经剑桥大学出版社许可重印]

这个幂律关系现在被称为克莱伯定律（Kleiber’s　law）。最近有研究发现，3/4次幂比例不仅对哺乳动物和鸟类成立，对鱼类、植物，甚至单细胞生物也成立。

克莱伯定律是建立在对代谢率和体重的测量的基础上，克莱伯没有解释这个定律为什么成立。结果克莱伯定律一直困扰着生物学家们。生命系统的重量范围很大，细菌不到万亿分之一克，鲸鱼则可能超过10万千克。这个规律不仅违背简单的几何推理，适用范围也大得惊人，涵盖各种大小的生物，也适用于各种生物类型和生境。到底是生物的哪种共性导致了这个简单而优雅的规律呢？

其他一些相关的比例关系也一直让生物学家们感到困惑。例如，越大的哺乳动物生命期越长。老鼠的生命期一般为2年左右，猪的生命期则大约为10年，大象超过50年。其中也有例外，特别是人类，但是对大部分哺乳动物都成立。如果画出许多物种的平均生命期和体重的关系，会发现是指数为1/4的幂律。如果画出平均心率与体重的关系，你会得到指数为-1/4的幂律（越大的动物心率越慢）。生物学家们发现了大量的幂律关系，都是分母为4的分数指数。因此，这些关系也被称为四分幂比例律（quarter-power scaling laws）。许多人怀疑，这些四分幂比例律意味着这些生物具有某种非常重要的共性。但是没人知道是什么共性。

一次跨学科合作

20世纪90年代中期，新墨西哥大学的生态学教授布朗（James Brown）多年来一直在研究四分幂比例律。他很早就意识到，如果能解决这个问题，理解这些普适比例律的原理，对于发展出生物学的一般理论将很重要。一位对比例问题很着迷的生物学研究生恩奎斯特（Brian Enquist）加入了布朗的团队，他们开始尝试一起来攻克这个问题（图17.3）。

布朗和恩奎斯特怀疑，向细胞输送营养的系统结构是解决这个问题的关键。血液不断在血管中循环，血管形成了一个树状网络，将营养物质输送到身体的所有细胞。同样，在肺部是由支气管组成的分支结构将氧气输送到血管提供给血液（图17.4）。布朗和恩奎斯特认为正是这种在动物体内普遍存在的分支结构导致了四分幂律。要理解这种结构为何会导致四分幂律，就得用数学描述这种结构，并从数学上证明这种结构直接导致了观察到的那些比例律。

大部分生物学家，包括布朗和恩奎斯特，都不具备必需的数学背景，无法进行这样复杂的几何和拓扑分析。因此布朗和恩奎斯特决定寻找一位“数学伙伴”——一位数学家或理论物理学家，可以帮助他们解决这个问题，同时又不过度简化以至于失去生物学意义。

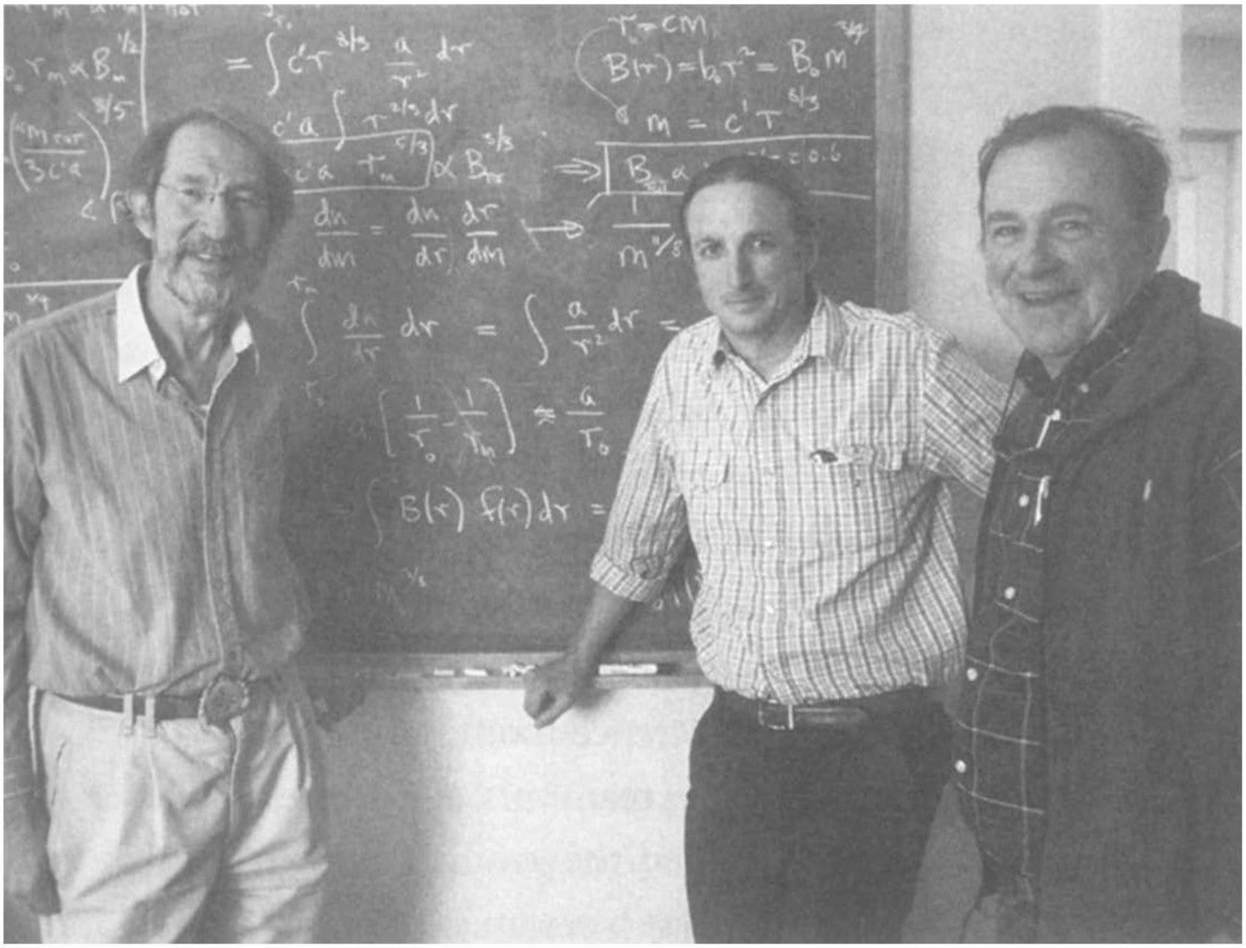


图17.3　从左往右：韦斯特、恩奎斯特、布朗（圣塔菲研究所拥有照片版权。经许可重印）

韦斯特（Geoffrey West）正符合他们的要求。韦斯特是理论物理学家，在洛斯阿拉莫斯国家实验室工作，他具有解决比例问题所需的数学能力。他研究过比例问题，虽然是在量子力学领域，但是他自己也思考过生物比例问题，只是不太懂生物学。20世纪90年代中期，布朗和恩奎斯特在圣塔菲研究所遇到了韦斯特。此后三人开始每周在研究所会面，形成了合作关系。我记得当时每星期都见到他们，在一间玻璃会议室里专心讨论，同时还有人（通常是韦斯特）在黑板上写下一堆复杂的公式。（恩奎斯特后来用“放烟火”来描述他们的数学结果。1）当时我并不是很清楚他们在讨论什么。但后来当我听到韦斯特在讲座上介绍他们的理论时，我被这个理论的优雅简洁和适用范围之广惊呆了。在我看来，这个工作是迄今为止复杂系统研究的巅峰之作。

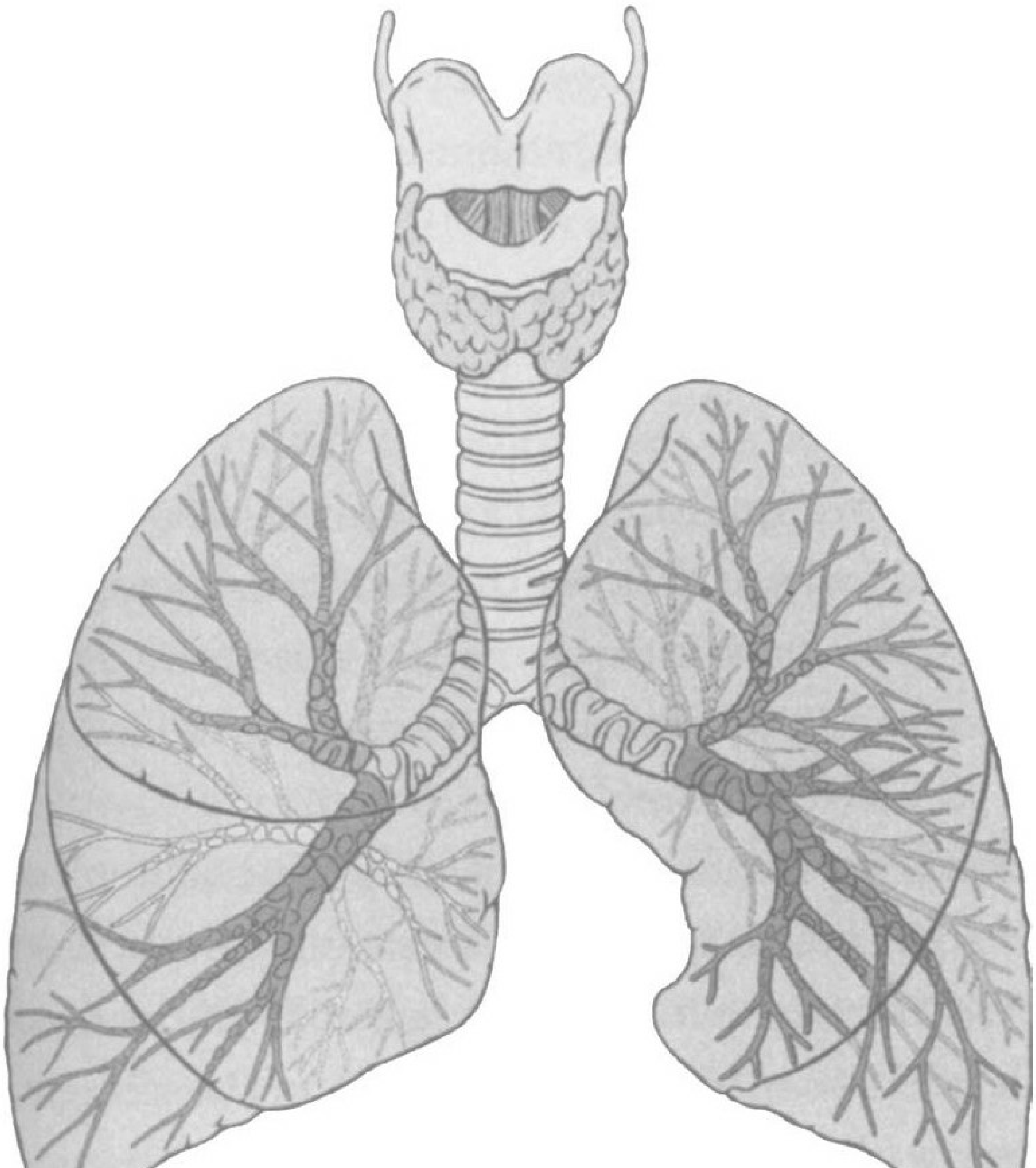


图17.4　肺中由支气管组成的分支结构（Patrick Lynch绘图，经知识共享组织许可使用，http：//creativecommons.org/licenses/by/3.0/）

布朗、恩奎斯特和韦斯特提出的理论不仅揭示了克莱伯定律和其他观察到的生物比例关系，而且还推断出生命系统中一系列新的比例关

系，其中许多后来都得到了数据支持。这个理论叫作代谢比例理论（Metabolic scaling theory，或简单代谢理论），结合了生物学和物理学，也在这两个领域引起了很大的轰动和争议。

幂律与分形

代谢比例理论回答了两个问题：①到底为什么代谢比例遵循幂律；②为什么遵循指数为3/4的幂律。在阐述这个理论之前，我需要简单阐释一下幂律与分形的关系。

还记不记得第7章讨论的科赫曲线和分形？如果记得，你也许还会记得“分形维”的概念。我们看到，在科赫曲线中，每一层的线段长度都是上一层的1/3，而每一层都是由上一层的4份拷贝组合而成。类似于维数的传统定义，我们这样定义科赫曲线的分形维：$3^{维数}=4$，得到维数=1.26。也就是说，如果每一层与上一层的比例系数是x，而又是由上一层的N个拷贝组合而成，那么$x^{维数}=N$。现在阅读了第15章之后，你会意识到这就是幂律，维数就是幂律指数。这说明幂律与分形有密切关联。我们在第15章图15.10中看到的幂律分布就是分形——它们在所有缩放尺度上都自相似，而幂律指数则是相应的分形维（参见第7章），维数量化的正是分布的自相似与放大倍数的比例关系。因此我们可以说，例如，网络的度分布具有分形结构，因为它是自相似的。同样，我们也可以说科赫曲线这样的分形导致了幂律——幂律描述的正是曲线的自相似与放大倍数的比例关系。

最后的结论就是，分形结构是产生幂律分布的一种方式；如果你发现某种量（例如代谢率）遵循幂律分布，你就可以猜想这是某种自相似或分形系统导致的。

代谢比例理论

考虑到代谢率是身体细胞将原料转化为能量的速率，布朗、恩奎斯特和韦斯特认为代谢率应当主要是由向细胞输送原料的效率决定的。而输送原料是生物循环系统的工作。

布朗、恩奎斯特和韦斯特意识到，对于循环系统来说，起决定性作用的不仅是它的质量或长度，更重要的是它的网络结构。正如韦斯特所

说的："你应当在两个不同的尺度上思考 [267] ——表面的你和真正的你，而后者是由网络组成的。"

在发展理论时，布朗、恩奎斯特和韦斯特假定进化过程已使得循环和营养输送系统能尽可能地填充身体空间，也就是说能将养分输送到身体所有部位的细胞。他们还假定进化出的网络能最小化向细胞输送养分时所花费的能量和时间。最后，他们假定网络向身体组织提供燃料的"终端单元"大小不随体重变化。这个看法是有根据的，例如，大多数动物循环系统的毛细血管都是一样大的。只是越大的动物毛细血管越多。这是因为细胞的大小不受身体大小影响：老鼠和河马的细胞都差不多大。只是河马的细胞更多，因此也需要更多的毛细血管向它们提供养分。

尽可能填充空间的几何对象其实就是分形分支结构——在所有尺度上自相似意味着空间在所有尺度上都被同等填充。布朗、恩奎斯特和韦斯特那些日子在玻璃会议室发展的精巧数学模型正是将循环系统视为填充空间的分形。他们融合了前面说的能量时间最小化和终端单位大小不变假设，然后问，当身体变大时，模型会发生什么变化呢？他们的计算表明，决定代谢率的养分输送速率与体重呈指数为3/4的比例关系。

推导出指数为3/4的模型细节相当复杂 [268] ，不过还是可以看看他们对指数3/4的解释。前面讨论了鲁伯纳的表皮假说，代谢率与体重的比例关系就如同表面积与体积的比例关系，指数为2/3。理解指数3/4的一种方式是将其视为表皮假说应用到四维生物的结果！通过简单的维数类比就能明白这一点。圆这样的二维对象有周长和面积。如果是三维，就分别对应表面积和体积。如果是四维，表面积和体积则分别对应于"表面"体积和超体积——这个量很难想象，因为我们天生擅长思考三维，不擅长思考四维。表面积与体积呈指数为2/3的比例关系，通过类似的论证，就可以知道四维的表面体积与超体积呈指数为3/4的比例关系。

简而言之，布朗、恩奎斯特和韦斯特的观点就是，进化将我们的循环系统塑造成了接近于"四维的"分形网络，从而使我们的新陈代谢更加高效。用他们自己的话说："虽然生物是三维的 [269] ，内部的生理结构和运作却表现为四维……分形几何给了生命一个额外的维度。"

理论的应用

代谢比例理论最初是用来解释各种动物的代谢比例，也就是图17.2中的现象。但是布朗、恩奎斯特和韦斯特，以及越来越多新加入的研究者并不满足于此。不断有新的生物门类和现象被发现可以用这个理论解释。他们认为他们的理论也可以解释其他四分幂比例律，例如心率、生命期、妊娠期以及睡眠时间等。他们认为植物的代谢比例也可以用这个理论解释，许多植物都是用类似于分形的管道网络输送水和养分。不仅如此，他们认为树干周长、植物生长速度等动植物特性的四分幂比例律都可以用这个理论解释。一种更广义的代谢比例理论将体温也囊括进来，可以解释爬行动物和鱼类的代谢律。

再来看看微观领域，研究组推测，他们的理论可以应用到细胞层面，3/4指数代谢比例既可以计算单细胞生物的代谢律，也适用于细胞内部分子层面的类代谢运输过程，甚至包括像线粒体这样的细胞器内部的类代谢过程。研究组还认为这个理论可以解释生物DNA的变化速率，因此与遗传学和进化生物学也密切相关。还有人发现这个理论可以解释肿瘤生长速度与重量的比例关系。

在大的方面，代谢比例理论及其扩展已经被应用到整个生态系统。布朗、恩奎斯特和韦斯特认为他们的理论能解释一些生态系统中观察到的种群密度与身体大小的-3/4幂律关系。

事实上，代谢对于生命系统的方方面面都很重要，几乎找不到这个理论没有触及的领域。许多科学家对此非常兴奋，不断为这个理论寻找新的应用。有人认为代谢比例理论“有统一整个生物学的潜力”[270]，“对于生物学的重要性就好比牛顿的发现对于物理学的重要性”[271]。研究组在一篇论文中这样说道：“可以预见，广义代谢理论的涌现对于生物学的重要性将类似于遗传理论。”[272]

争议

一个刚刚初具雏形的理论却声称能解释这么多东西，可想而知，一些科学家会对代谢比例理论充满热情，一些人则极力反对。下面是最近发表在顶级科学期刊上的两个主要批评观点：

◆四分幂比例律并不像理论说的那样具有普适性。通常来说，任何关于生物系统的一般性属性都会有特例。（甚至规则本身都会有特例。）姑且认为代谢比例理论也不例外。虽然大部分生物学家都同意大多数物种都遵循各种四分幂比例律，但也有许多例外。甚至在单一物种内的代谢律都变化很大。狗就是一个常见的例子，小型犬与大型犬的寿命都差不多。有观点认为，克莱伯定律只是统计平均，偏差有可能相当大，而代谢理论无法解释这一点，因为只考虑了体重和体温。有人则认为理论得出的一些规律与实际数据严重不符。甚至有人认为克莱伯根本就是错的 [273] ，一百多年前鲁伯纳提出的表皮假说才是对的，指数为2/3的幂律对数据的拟合最好。在大部分情形中，争议都在于对代谢比例的数据应该如何解读，以及“拟合”指的是什么。代谢比例研究组坚持他们的理论，并且不厌其烦地回应了许多争议，由于涉及高等统计学和生物学，这些争议变得越来越艰深晦涩。

◆克莱伯的比例定律是对的，但代谢比例理论错了。有些人认为代谢比例理论过度简化，生命极为复杂多变，不可能被单一理论所涵盖，分形结构也不是解释幂律分布现象的唯一途径。一位生态学家这样评论：“人们对于所涉及的生理细节了解得越多 [274] ，这种解释就越显得不合理。”另一位学者则说：“事情简单当然很好 [275] ，但现实中往往不是这样。”另外还有人认为代谢比例理论的数学有错误 [276] 。代谢比例理论研究组坚决不同意这一点，并且指出了批评意见中的一些基本数学错误。

研究组坚持自己的立场，他们对吹毛求疵的批评意见感到沮丧。[277] 韦斯特说：“我的内心不会向这些在我脚边乱吠的小狗屈服 [278] 。”不过研究组还是认为，有这么多批评意见不是坏事——不管他们到底怎么认为，毕竟有很多人在关注代谢比例理论。并且，就像我在前面提到的，怀疑是科学家们最重要的职责，越是杰出而有雄心的理论，越是会受到质疑。

争议不会很快平息；牛顿的引力理论提出来60年后都没有被广泛认可，许多最重要的科学进展都曾有类似的经历。现在我们能说的是，代谢比例理论非常有趣，应用范围很广，也得到了一些实验数据的支持。生态学家穆勒—兰道（Helene Müller-Landau）评论道：“我想韦斯特和恩奎斯特等人不会一直重复他们的观点， [279] 批评者也不会一直重复他们的质疑，随着时间流逝，证据的天平最终会倒向胜利的一边。”

幂律的未解之谜

在前一章和这一章我们看到了许多幂律。除了这些，在城市规模、收入、地震、心率变化、森林火灾和股市波动等现象中都发现了幂律分布，这还只是其中一小部分。

第15章曾讲过，科学家们一般都假定大部分自然现象都服从钟形曲线或者说正态分布。然而幂律却在很多现象中都有被发现，以至于一些科学家说它“比‘正态’还要正态”。[280] 用数学家维林格（Walter Willinger）和他同事的话说：“在复杂的自然和工程系统中获得的数据中发现（幂律）分布，应当视为正常而不是意外。”

科学家们对自然界中钟形曲线分布的成因有很好的理解，但幂律在一定程度上却还是个谜。我们已经看到，对于自然界中观察到的幂律有各种解释（例如，偏好附连、分形结构、自组织临界性、高度容错等等），对于是哪种机制导致了幂律现象很少有共识。

20世纪30年代早期，哈佛语言学教授齐普夫（George Kingsley Zipf）在一本书中介绍了语言的许多有趣属性。随意在小说或报纸中取一大段文字，将所有词根据出现次数排序。例如，下面是莎士比亚戏剧《哈姆雷特》的独白“生存还是毁灭”中的词频表：

词	频率	排序
the	22	1
to	15	2
of	15	3

and	12	4
that	7	5
a	5	6
sleep	5	7
we	4	8
be	3	9
us	3	10
bear	3	11
with	3	12
is	3	13
’tis	2	14
death	2	15
die	2	16
in	2	17
have	2	18
make	2	19
have	2	18
end	2	20

根据词频降序排列，频数最高的词排第一（“the”），频数第二高的词排第二，等等。一些词的频数一样（例如，“a”和“sleep”都出现了5次），对于这种情况随机排序。

在图17.5中画出了“生存还是毁灭”的词频与排名的关系。图的形状接近幂律。如果选取的文本更多，图形会更接近幂律。

齐普夫用这种方法分析了大量文本（没有借助计算机），他发现，对于大规模文本，词频大致正比于其排名的倒数（也就是1/排名）。这是指数为-1的幂律。排名第二的词的频数大约是排第一的词的一半，第三大约是1/3，等等。这个关系现在被称为齐普夫定律[281]（Zipf’s law），这可能是最著名的幂律。

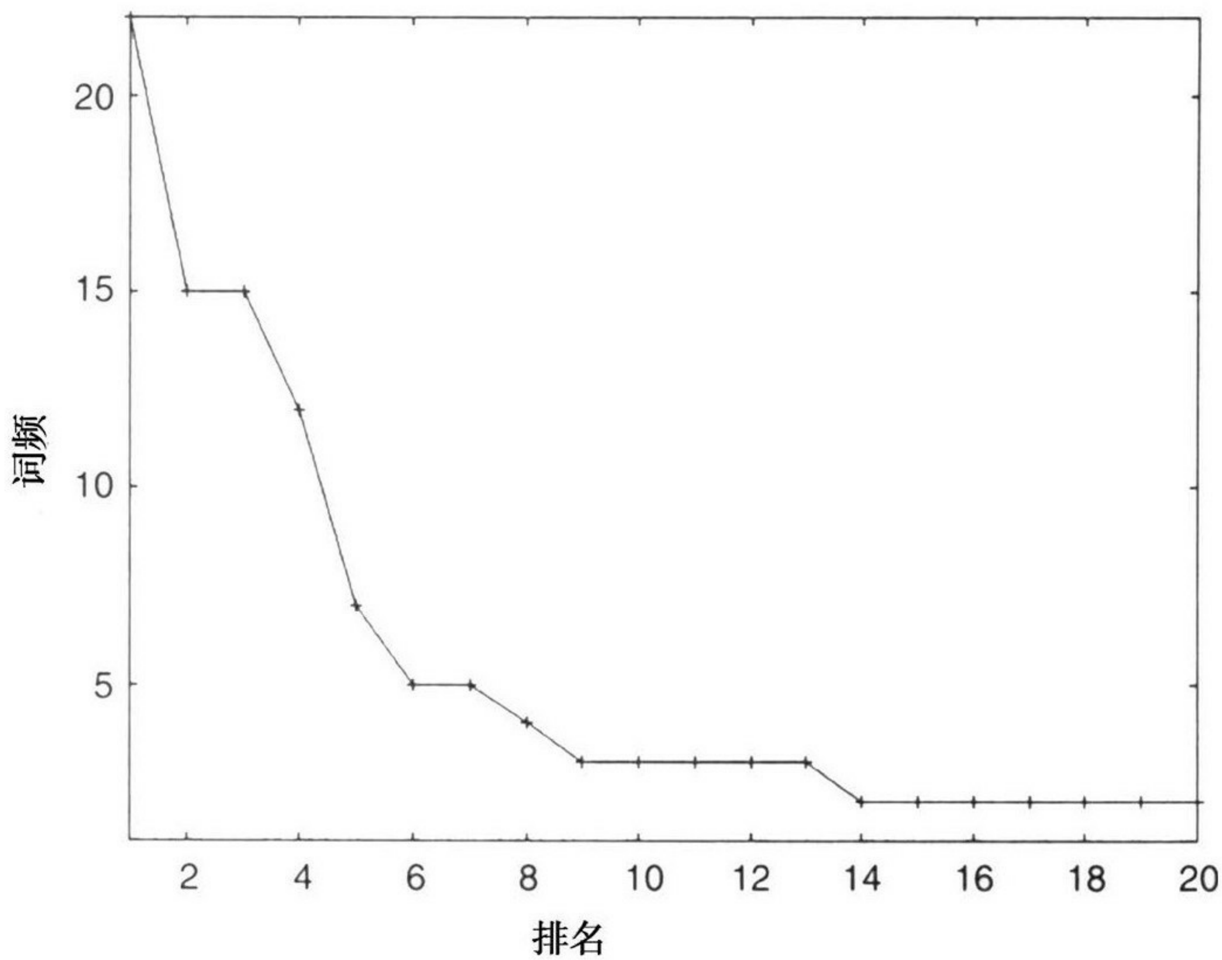


图17.5　齐普夫定律，以莎士比亚的独白“生存还是毁灭”为例

对齐普夫定律有各种解释，齐普夫自己提出，一方面，人们一般都遵循“最省力原则（Principle of Least Effort）”：一旦用到了某个词，对类似的意思再用这个词就比换其他词要省力。另一方面，人们希望语言没有歧义，这又需要用不同的词来表示相似却又不完全一样的意思。齐普夫从数学上证明了这两种倾向在一起会产生观察到的幂律分布。

20世纪50年代，因发现分形而闻名的曼德布罗特从信息量的角度提出了不同的解释[282]。借鉴香农的信息论（参见第3章），曼德布罗特将词视为“讯息”，发送者在将信息量最大化的同时尽量将发送信息的成本最小化。例如，“feline”和“cat”的意思都是猫，但后者更短，因此传送成本也更低（或者更节省能量）。曼德布罗特证明，如果同时优化信息量和传送成本，就会导致齐普夫定律。

几乎同时，西蒙（Herbert Simon）也提出了一种解释[283]，可以说是偏好附连的前身。西蒙设想一个人每次向文本中添加一个词。他认为，人们重用一个词的概率正比于这个词在文本中的当前频数。没出现过的词具有同样的非零概率。西蒙证明这个过程产生的文本会遵循齐普夫定律。

对于曼德布罗特和西蒙的解释哪个正确，争论很激烈[284]（从《信息与控制》杂志不断收到的信件可见一斑）。

结果，几乎与此同时，让所有人都大跌眼镜，心理学家乔治·米勒（George Miller）使用简单的概率论证明，[285]让猴子在键盘上随意敲击，如果（偶然）敲到了空格键就断词，这样得出的文本同样遵循齐普夫定律。

20世纪30—50年代针对齐普夫定律提出的许多解释是目前针对自然界中产生幂律的物理或信息机制的争论的缩影。理解幂律分布的根源、意义和在各学科中的共性，是目前许多复杂系统研究领域最为重要的未解决的问题。我相信，随着这些现象背后的科学越来越清晰，你还会不断听到这个问题的消息。

第18章　进化，复杂化

在第1章我曾问过："生物进化是如何产生出个体如此简单、整体上又如此复杂的生物呢？"通过书中的例子我们可以看到，对生命系统的理解越深入，就越感到惊讶，这样精巧的复杂性居然是通过有利突变和历史偶然的逐步积累形成的。这也正是从达尔文时代到现在神创论或其他超自然"智能设计"的拥护者论证的依据。

进化是如何创造出复杂性，或者说是否能创造复杂性，以及生物复杂性该如何刻画和度量，这些问题都还没有解决。复杂系统研究几十年来最重要的贡献之一就是为这些老问题提供了一些新的研究途径。这一章我将介绍遗传学和基因调控动力学的最新发现，它们为我们带来了一些关于复杂系统进化的惊人的新认识。

遗传，复杂化

在科学研究中，经常有一些新的技术会打开新发现的闸门，从而改变以前建立起来的对研究领域的认识。回到第2章我们可以看到一个这样的例子——计算机的发明使得为天气这样的复杂系统建模仿真成为可能，并因此揭示了混沌的存在。最近，超级天文望远镜的建造在天文学领域导致了关于所谓的暗物质和暗能量的许多新发现，因此引发了对之前的宇宙学知识的重新审视。

过去四十年中，没有什么技术的影响比得上所谓的分子革命对遗传学的影响。对DNA进行快速复制、测序、合成，实施DNA工程，对分子层面的结构进行成像，同时观察数以千计不同基因的表型，这些技术还只是20世纪末21世纪初生物技术取得成就的一小部分。随着新技术的出现，生物学家们可以更细致地观察细胞，更多出人意料的复杂性也随之出现。

在沃森和克里克发现DNA结构之后，DNA基本被视为由基因组成的序列，每个基因编码一种特定的蛋白质，在细胞中实现一定的功能。基因序列本质上被视为细胞的"计算机程序"，通过RNA、核糖体等物质的译码和执行，合成出相应的蛋白质。DNA在复制过程中会有小的随机变化；对有利变化的长期积累最终会导致生物的适应性变化，并产生新的物种。

这种传统观念在过去40年中已经发生了巨变。分子革命一词不仅指

遗传学中的新技术，也指这些技术带来的对DNA、基因和进化本质的革命性新观点。

基因是什么

分子革命的一个诱因就是基因概念本身。第6章描述的DNA的机制仍然是成立的——染色体中包含DNA，通过转录和译码产生蛋白质——但这只是故事的一部分。下面来看看部分新近发现的许多现象，这些现象关注的是基因和遗传的运作机制。

◆基因并不像“一根绳子上串着的豆子”。我在中学学生物时，基因和染色体被解释比喻成一根绳子上串着的豆子（我记得我们还用塑料豆子组装过模型）。后来发现基因并不是相互分开的。有些基因相互重叠——也就是说，它们各自编码不同的蛋白质，但是共用DNA核苷酸。有些基因甚至完全包含在其他基因内部。

◆基因可以在染色体上移动，甚至移动到其他染色体。你也许听说过“跳跃基因（jumping genes），”实际上基因是可以移动的，染色体的组成也会被重新排列。这在任何细胞中都有可能发生，包括精子和卵子，也就是说可以遗传。这样产生的变异率比DNA复制错误导致的变异率要高得多。一些科学家提出，近亲甚至同卵双胞胎之间的差别可能就是这种“可动遗传因子[286]（mobile genetic element）”造成的。还有人提出，跳跃基因是导致生命多样性的机制之一。

◆单个基因可以编码多个蛋白质。以前一直以为基因和蛋白质是一对一的关系。这个认识在人类基因组被测序后受到怀疑，基因编码的不同蛋白质的数量可能超过100000种，而人类基因组只有大约25000个基因。最近发现的多重剪接（alternative splicing）和RNA编辑（RNA editing）可以帮助解释这个差异。这些过程可以在信使RNA 转录DNA之后和译码成氨基酸之前以各种方式变化。这意味着同样的基因通过不同的转录事件可以产生出不同的蛋白质。

◆由于如此复杂，以至于最专业的生物学家也经常无法对“基因”的定义达成共识。最近一组科学哲学家和生物学家进行了一项调查[287]，向500名生物学家各提供一些不常见但真实的DNA序列，然后问他们这些序列是不是“基因”，以及他们对自己的答案有多大把握。结果发现对

其中许多序列，他们的想法产生了分化，60%的人相信一个答案，40%的人相信另一个答案。《自然》杂志上报告这项调查的文章评论道："对分子遗传学越专长的学者[288]，越不确定基因到底是什么。"

◆生物系统的复杂性主要来自基因网络，而不是单个基因独立作用的简单加总。第16章曾讲过，基因调控网络目前是遗传学的研究重点。以前的绳子串豆子的观念同孟德尔遗传律一样，都是把基因看作线性的——每个基因都各自负责某个表型。而现在的普遍观念则是，细胞中的基因组成了非线性的信息处理网络，一些基因会根据细胞状态控制其他基因的行为——基因并不是独立运作。

◆即使基因的DNA序列不发生变化，基因的功能也会发生可遗传的变化。最近兴起的表观遗传学（epigenetics）研究的就是这种变化。一个例子就是所谓的DNA甲基化（methylation），细胞中的一种酶将特定的分子连接到DNA序列的某些部分，将这些部分"关闭"。一旦细胞中发生这种现象，这个细胞的所有后代就会产生同样的DNA甲基化。如果DNA甲基化发生在精子或卵子中，就会被遗传。

◆一方面，这类表观遗传现象在所有细胞中都不断在发生，对生命活动的许多方面都很关键，因为它可以关闭不再需要的基因（例如，一旦进入成年期，我们就不再需要像小孩一样生长发育，控制青春期发育的基因就会甲基化）。另一方面，错误的甲基化，或者应当甲基化却没有甲基化，又会导致遗传紊乱和疾病。事实上，一些人认为，正是由于胚胎发育期缺乏必需的甲基化[289]，使得很多克隆胚胎无法存活，许多克隆动物即便存活也会有严重甚至致命的缺陷。

◆最近发现，在大部分生物中，DNA转录为RNA之后很大部分最终都没有被译码成蛋白质[290]。这些所谓的非编码RNA对基因和细胞的功能具有调控作用，这些以前都认为是由蛋白质单独完成的。非编码RNA的作用是目前遗传学中一个非常活跃的研究领域[291]。

遗传学已经变得非常复杂了。这种复杂对生物学的影响巨大。2003年，人类基因组计划发布了完整的人类基因组——人类DNA的全部序列。虽然这个计划获得了大量新发现，但还是没有达到许多人的预期。一些人曾以为人类基因的详尽图谱能让我们彻底理解遗传的运作原理，哪个基因对应哪项特征，并带来医学发现和靶向性基因治疗的革命。虽

然发现了一些基因可能是某些疾病的原因，但结果表明仅仅知道DNA的序列还不足以让我们理解人（或其他复杂生物）的全部特性和缺陷。

使得基因序列被寄予如此厚望的一个主要因素是国际生物技术工业。《纽约时报》最近的一篇文章报道了新近发现的这些遗传复杂性对生物技术工业的影响："基因独立运作的想法是1976年之后形成的[292]，这也是第一家生物技术公司成立的时间。事实上，整个生物技术工业的经济基础都建立在这个认识之上。"

问题还不仅仅在于遗传学被迅速修改。生物技术工业一个潜藏的大问题是基因专利的归属。几十年来，生物技术公司对人类DNA中一些被认为"编码了特定的功能性产物"[293]的序列申请了专利。但就像我们在前面看到的，许多复杂特性并不是由某个基因的DNA序列单独决定的。既然这样，这些专利还能成立吗？如果"功能性产物"是表观遗传过程作用或调控基因的结果呢？又或者这个产物不仅需要被申请专利的基因，还需要调控这个基因的基因，以及调控调控基因的基因呢？如果这些调控基因的专利被授予了其他人呢？一旦放弃线性基因的观念，面对本质上的非线性，这些专利的意义就会变得不明不白，好处是专利律师和法官今后不用担心失业了。专利不是唯一的问题。就像《纽约时报》指出的："基因组网络化的证据实际上毁掉了对当今生物技术产品商业化进行的所有官方风险评估的科学基础[294]，不管是转基因作物还是医药。"

不仅遗传学，整个进化论都因这些新的遗传学发现受到了挑战。"进化发育生物学（evolutionary developmental biology）"领域就是一个突出的例子。

进化发育生物学

进化发育生物学是一个让人兴奋的领域，这个领域最近的发现据称解释了至少3个遗传和进化的大谜团：

1.人类只有大约25000个基因。复杂性从何而来？

2.人类在遗传上与其他许多物种很类似。例如，我们的DNA超过90%与老鼠一样[295]，超过95%与大猩猩一样。为什么我们的形态与这

些动物相差这么大？

3.如果古尔德等人提出的进化间断平衡是正确的，身体形态为何会在很短的进化时期内发生巨大变化呢？

最近有观点认为，这些问题的答案至少部分在于基因开关（genetic switch）的发现。

发育生物学和胚胎学研究的是单个受精卵如何变成由亿万细胞组成的活生物体的过程。然而，现代综合关注的却是基因；用发育生物学家卡罗尔（Sean Carroll）的话说，现代综合是将胚胎和发育过程视为“‘黑箱’，在其中可以将遗传信息变成三维的、活的动物”[296]。这部分归咎于以前的观念，认为多种多样的动物形态最终可以用基因的数量和DNA构成的巨大差异解释。

20世纪80—90年代，这种观念有了很大变化。就像前面说的，DNA测序发现许多不同的物种DNA却很相似。遗传学的进展也使得对胚胎发育过程中基因表达的机制有了更详细的了解。结果发现这些机制与之前预想的差别很大。胚胎学家发现，在研究过的复杂动物中，都存在一小部分“主导基因”调控动物许多身体部位的发育成形。更让人吃惊的是，各物种之间，不管是果蝇还是人类，虽然形态差异极大，主导基因的DNA序列却有许多是相同的。

如果发育过程是受相同的基因掌控，这些动物的形态怎么会如此不同呢？进化发育生物学的支持者提出，物种形态多样性的主要来源不是基因，而是打开和关闭基因的基因开关。这些开关是不编码蛋白质的DNA序列，通常长度为几百个碱基对。它们以前被认为是所谓的“垃圾基因”的一部分，但现在发现有基因调控的作用。

图18.1是基因开关的工作原理。基因开关是位于某个基因旁边的非编码DNA序列。这个序列一般包含有一组标签子序列，其中每个都可以与特定的蛋白质结合，从而让蛋白质附着到DNA上。旁边的基因是否能被转录，速度多快，都取决于蛋白质与这些子序列的结合情况。允许转录的蛋白质会为进行转录的RNA分子创造强结合点；阻止转录的蛋白质则会阻挡这些RNA分子同DNA结合。其中一些蛋白质还有能消除其他蛋白质的作用。

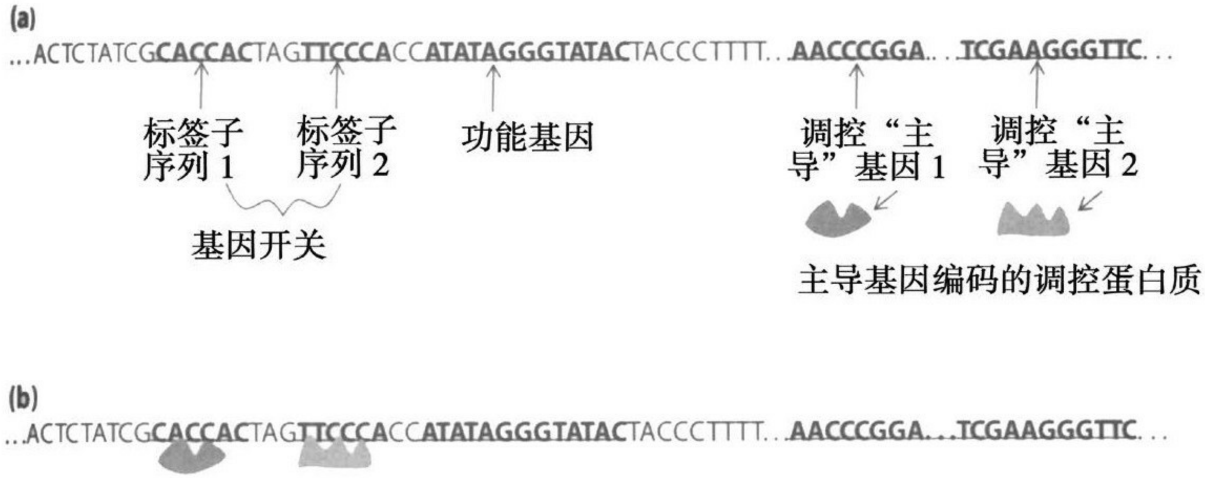


图18.1　基因“开关”示意图。（a）一段DNA序列，包含一个有两个标签子序列的开关、被开关打开的功能基因以及两个调控主导基因。调控主导基因生成调控蛋白质。（b）调控蛋白质同标签子序列结合，打开功能基因——也就是允许其转录

这些特定的调控蛋白质是从何而来呢？同所有蛋白质一样，它们由基因生成，这里是由调控基因对这些蛋白质进行编码，根据细胞的当前状态决定相应基因是开还是关。这些调控基因又是如何判断细胞的状态呢？一些蛋白质可以与这些调控基因本身的开关相结合，从而向其传递细胞状态的信号。这些蛋白质通常是由另外的调控基因编码，这些基因又由其他基因调控。

简而言之，基因调控网络包括功能基因和调控基因，功能基因编码用于细胞结构和运转的蛋白质（和非编码RNA），而调控基因编码的蛋白质则可与目标基因旁边的DNA“开关”相结合，从而开启或关闭相应的基因。

现在可以为这节开头提出的3个问题给出进化发育生物学的答案了。人类（和其他动物）比基因数量所表现的要复杂得多，可能的原因很多，“基因是什么”一节列出了一些。但主要原因还是基因调控网络使得基因表型的可能性极多，因为蛋白质对基因开关的附着情况有很多种。

人类之所以与其他差别极大的物种能有如此多相同的基因，是因为虽然基因是一样的，但基因开关的序列构成却已进化得不一样了。基因开关的微小变化能导致发育过程中基因采取截然不同的开关模式。因此，根据进化发育生物学，生物的多样性主要来自开关而不是基因的进

化。这也是为什么形态的巨大变化——可能还包括物种形成——可以在很短的进化时间内发生：主导基因不变，但是开关变了。根据进化发育生物学的观点，进化的主要力量正是这种——长期以来一直被视为“垃圾”的DNA的——变化，而不是新基因的出现。生物学家马蒂克就此评论说：“讽刺的是……一直被视为垃圾的[DNA]却藏有人类复杂性的秘密。”[297]

进化发育生物学的一个惊人例证就是燕雀鸟喙的进化。第5章曾讲过，达尔文发现加拉帕格斯群岛燕雀的喙的大小和形状差别很大。直到不久前，大部分进化生物学家都还认为这种差别是几种基因随机变异逐渐积累的渐变过程。但最近发现了一个名为BMP4的基因可以通过调控生成骨骼的基因来控制喙的大小和形状。鸟在发育过程中BMP4的表达越强烈，喙就越强大。另一种名为钙调素（calmodulin）的基因则被发现与长细形的喙有关。卡罗尔·尹（Carol Kaesuk Yoon）在《纽约时报》上撰文介绍，“为了证明BMP4基因确实能触发生长粗壮、能打开坚果的喙，[298] 研究者在小鸡胚胎发育出喙时人为加快了BMP4的产生。结果小鸡长出了宽厚而结实的喙，类似于能啄开坚果的燕雀……像BMP4一样，钙调素基因的表达越强，雀喙就会长得越长。如果在小鸡胚胎中人为增加钙调素，小鸡就会长出变长的喙，就像啄食仙人掌的燕雀……这样科学家就发现，无需几十上百种基因，只需这两种，就有可能让鸟喙变得或是厚重，或是短粗，或是细长”。结论是鸟喙（及其他特征）形态的巨大变化可以很快发生，而无须等待时间漫长的随机变异。

进化发育生物学挑战进化传统观念的另一个例子是趋同进化（convergent evolution）。在中学生物课上我们学过，章鱼眼睛和人类眼睛——形态差异很大——是趋同进化的例子：这两个物种的眼睛是相互独立进化出来的，是自然选择作用于两种不同环境的产物，两种环境中眼睛都具有适应优势。

然而，最近有证据表明，这两种眼睛的进化并不像以前认为的那样独立。人类、章鱼、苍蝇等物种都具有名为PAX6的基因，这种基因能引导眼睛的发育。瑞士生物学家格林（Walter Gehring）做了一个古怪而富有启发的实验，[299] 实验中格林将老鼠的PAX6基因取出插入到果蝇的染色体中。在不同实验中，PAX6被插入染色体的三个不同部位：这三个部位分别引导腿、翅膀和触须的发育。结果非常怪异：果蝇的

腿、翅膀和触须上长出了类似眼的结构。这种结构像果蝇的眼，而不是老鼠的眼。格林得出结论：眼睛不是多次独立进化出来的，而是只有一次，有一个具有PAX6基因的共同祖先。这个结论在进化生物学家中仍然极具争议。[300]

虽然主导基因引导的基因调控网络能产生多样性，它们也对进化施加了一些限制。进化发育生物学家认为任何生物的身体形态类型都受主导基因高度约束，这也是为什么自然界中只有少数基本的身体结构类型。如果基因组很不相同的话，也许会有新的身体结构类型，但实际上进化无法让我们变成那样，因为我们非常依赖现在的调控基因。我们的进化可能性是有局限的。根据进化发育生物学的观点，“所有特性都能无限变化”的观念是错误的。

基因调控和考夫曼的“秩序的起源”

理论生物学家考夫曼（图18.2）早在40年前就开始思考基因调控网络及其对进化的影响，当时进化发育生物学还尚未发端。同时他还思考了从这种复杂网络中涌现出的秩序对进化的意义。



图18.2　考夫曼（Daryl Black摄影，经许可重印）

考夫曼是复杂系统的传奇人物。我第一次遇见他是在我读研究生最后一年参加的一次会议上。他的演讲被安排在会议的最开头，我必须说，当时对我来说，那是我听过的最富启发的演讲。我不记得主题具体是什么；我只记得当时听演讲的感觉，我感到他讲得非常深刻，他提出的问题极为重要，我想去和这家伙一起做研究。

考夫曼刚开始是研究物理学，但很快就转向了遗传学研究。他的研究新颖而且极具影响，为他赢得了许多学术荣誉，包括麦克阿瑟“天才”奖，以及圣塔菲研究所的职位。在SFI的研讨会上，考夫曼经常会在听众席上插话：“我知道我只是个孤陋寡闻的乡下博士，但……”然后针对他以前不知道的一些很专业的主题滔滔不绝地讲述他即兴的想法，他可以讲上5分钟或更长时间。一位科学记者称他为“世界级的知识即兴演奏家”[301]，这个评价我认为相当中肯。

考夫曼说自己是“孤陋寡闻的乡下博士”只是谦虚。他是最有影响力的复杂系统思想家之一，富有远见，绝不是什么孤陋寡闻的人物。SFI有个笑话，说考夫曼“拥有达尔文进化论的专利”。事实上，他确实拥有在实验室演化蛋白质序列的专利技术，这项技术可以用来发现新的药物。

随机布尔网络

考夫曼可能是第一个发明和研究基因调控网络的简化计算机模型的人。他的模型结构是所谓的随机布尔网络（Random Boolean Network, RBN），是从元胞自动机扩展而来。同其他网络一样，RBN也包含节点以及节点之间的边。类似于元胞自动机，RBN以离散时间步的形式更新节点状态。在每一步各节点可以处于开状态或关状态。

状态只有开和关两种，这也是布尔一词的由来：布尔规则（或函数）的输入是一些等于0或1的数，根据输入得到的输出也要么是0要么是1。这类规则因数学家布尔而得名，布尔对其进行了深入的研究。

在RBN中，边是有方向的：节点A有边指向节点B，节点B并不必然（但可能）有边指向节点A。各节点的入度（从其他节点指向这个节点的边的数量）都是一样的，记为数字K。

下面来看看如何构建一个RBN：对每个节点，随机选取K个节点（可能包括自己），建立从这些节点指向目标节点的边，然后给这个节点随机赋予一条布尔规则，规则的输入为K个开关状态，输出则为单个开关状态[图18.3（a）]。

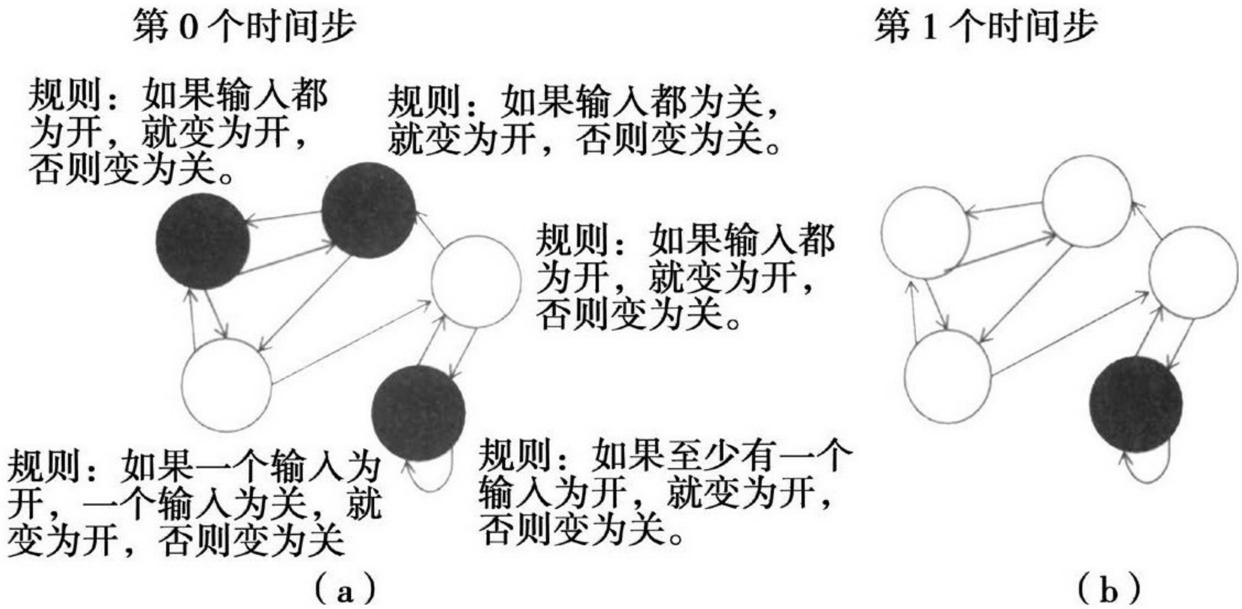


图18.3　（a）5个节点组成的随机布尔网络。各节点的入度（K）等于2。在第0个时间步，各节点处于随机的初始状态：开（黑）或关（白）。（b）各节点更新状态后第1个时间步的网络

RBN运行前，给每个节点赋予随机选择的初始开关状态。在每个时间步，各节点将自己的状态传送到其连接的节点，并从连接到它的节点接收状态输入。然后各节点根据其规则和输入确定自己下一步的状态。图18.3显示了一个5节点的RBN在第1步的变化，每个节点有2个输入。

RBN类似于元胞自动机，但是有两个主要区别：节点不是与空间上相邻的节点相连，而是随机连接，另外元胞自动机各节点的规则是一样的，而RBN的每个节点都有自己的规则。在考夫曼的研究中，RBN被作为基因调控网络的理想模型，其中节点代表“基因”，节点A指向节点B的边则表示“基因A调控基因B”。这个模型显然比真正的基因网络要简单得多。现在在生物学中使用这种理想模型已经很普遍，但在20世纪60年代考夫曼开始研究时还很少有人接受。

混沌边缘的生命

考夫曼和他的学生及同事用各种入度值K进行了大量RBN实验，从

随机初始状态开始，然后迭代许多步，节点状态一开始会以类似随机的方式变化，但最终要么停在不动点上（所有节点状态保持不变），要么进入周期振荡（整个网络的状态以较短的周期振荡），要么就根本不停下来，迭代了很久仍然像是随机变化。这种变化实际上是混沌，因为网络状态变化的轨迹敏感依赖于网络的初始状态。

考夫曼发现最终的行为是由网络中的节点数量以及各节点的入度K决定的。随着K从1开始（各节点只有一个输入）逐渐增大到等于全部节点的数量（每个节点都从其他所有节点有输入，包括它自己），RBN的典型行为经历了3种类型（不动点、振荡、混沌）。你可能注意到了这与逻辑斯蒂映射随着R增大时的行为变化类似（参见第2章）。在K=2时考夫曼发现了一个有趣的类型——既不是不动点、振荡，也不是完全混沌。类似于逻辑斯蒂映射的“混沌的发端”，他称之为“混沌边缘”。

考夫曼认为RBN能反映真实的基因网络的行为，他用水在各种温度下的状态作类比：“控制从受精卵到成人的发育的基因组网络 [302] 有可能处于以下三种类型：类似冰冻的有序态，类似气体的混沌态，以及位于有序和混沌之间类似液态的区域。”

考夫曼认为，生物既要具有活性同时又要稳定，RBN所模拟的基因网络应当为“液态”才有意义——既不能太僵硬也不能太混沌。用他的话说：“生命存在于混沌的边缘。”[303]

考夫曼用动力系统理论的术语——吸引子、分叉、混沌——来阐述他的发现。将节点可能的状态组合称为网络的全局状态。由于RBN的节点数量有限，因此可能的全局状态的数量也是有限的。如果网络迭代足够长时间，就必然会重复之前出现过的全局状态，接下来就会循环。考夫曼将这种循环称为网络的“吸引子”。通过反复实验，他估计出K=2时产生的不同吸引子的平均数量大致等于节点数量的平方根。

然后考夫曼对模型进行了大胆的解读。身体中所有细胞都具有同样的DNA。然而，细胞的类型却有很多：皮肤细胞、肝细胞，等等。考夫曼认为确定具体细胞类型的就是细胞中的基因表达形式——前面我曾讲过不同细胞的基因表达形式可以差别很大。在RBN模型中，一种吸引子就对应一种“基因表达”形式。因此考夫曼提出他的网络中的吸引子就对应于生物体内的一种细胞类型，细胞类型的数量应当接近100000的平方

根，大约316种。这与在人类身上发现的细胞种类数量——约为256种——相差不大。

在考夫曼进行这些计算的时候，普遍认为人类基因组包含大约100000个基因（因为人体中大约有100000种蛋白质）。考夫曼很兴奋，他的模型很好地预测了人体中细胞种类的数量。现在我们知道人类基因组只有25000个基因，因此根据考夫曼的模型得出的预测是大约158种。

有序的起源

模型并不完美，但考夫曼相信它揭示了他对生命系统最重要的一般性观点：原则上自然选择对于复杂生物的产生并不是必需的。许多K=2的RBN也能表现出“复杂”行为，而其中并不涉及自然选择或进化算法。他的观点是，一旦网络结构变得足够复杂——即有大量节点控制其他节点——复杂和“自组织”行为就会涌现出来。他说：

> 继承了达尔文传统的大部分生物学家都认为1，个体发育的有序性有赖于某种分子级的鲁布·戈德堡机器慢慢打磨2，这种机器是通过进化一点一滴累积而成的。我的观点相反：个体发育过程中的美丽秩序大部分是自发的，是极度复杂的调控网络所包含的惊人自组织的自然表达。我们似乎从根本上就错了。秩序，无处不在而且有生命力的秩序，是自然发生的。

考夫曼深受统计力学框架的影响。第3章曾介绍过统计力学。记得统计力学解释了温度这样的属性如何产生自对大量分子的统计。也就是说，人们不用跟踪所有分子的牛顿轨道就能预测系统的温度变化。考夫曼认为他的发现类似于统计力学的定律，其决定的是从大量相互连接调控的组分中涌现出来的复杂性。他称这条定律为“候选的热力学第四定律”[304]。热力学第二定律指出宇宙有熵增的内在趋势，而考夫曼的“第四定律”则提出生命具有复杂化的内在趋势，而这独立于自然选择的任何趋势。这个思想在考夫曼的《秩序的起源》（The Origins of Order）一书[305]中有详尽阐释。考夫曼认为，复杂生物的进化部分是由于这种自组织，部分由于自然选择，而且可能自组织才是起主导作用的，对选择的可能性施加了严格限制。

对考夫曼的研究的反响

由于考夫曼的研究意味着“对选择在进化论中的地位进行根本性的重新解读”[306]，可想而知人们对其的反应会有多强烈。这项研究有许多狂热的拥护者（“他的方法开启了新的未来”[307]；这是“对整个生物学进行建模的第一次严肃尝试”[308]）。另一方面，许多人则极为怀疑他的结果和他对结果的宽泛解释。一位评论家说考夫曼的写作风格具有“危险的诱惑力”[309]，并且这样评价《秩序的起源》的：“有时候就好像在虚幻的空间悠然漫步[310]，似乎无需理会来自现实的琐屑烦恼。”

实际上，相关的实验证据并不都是站在考夫曼一边。考夫曼自己也承认，将RBN作为基因调控网络的模型需要许多不符合实际的假设：每个节点只能两个状态取其一（而基因表达则有各种程度的强度），每个基因的调控基因都一样多，所有节点都是以离散时间步同步更新。这些简化可能会忽略基因活动的重要细节。

更麻烦的是，“噪声”——错误和其他不确定行为的来源——在真实世界的复杂系统中是不可避免的，基因调控也是如此。生物的基因网络会不断犯错误，但是它们具有弹性——大多数情况下我们的健康不会受这些错误影响。然而，仿真结果表明，噪声对RBN的行为有明显的影响[311]，有时候甚至会阻止RBN到达稳定的吸引子。虽然考夫曼对RBN的实验结果做了具体说明，但其中一些也经不起仔细推敲。

例如，考夫曼认为一个典型的网络可能的吸引子数量大约为节点数量的平方根，而他将这个解释为细胞类型。但进一步的仿真表明，吸引子的数量实际上并不是近似等于节点数量的平方根。当然，这并不意味着考夫曼的观点在大的方面就肯定错了；这只是表明，在发展更准确的模型方面还有很多工作要做。发展基因调控网络的准确模型现在是生物学中非常活跃的研究领域。

总结

生命系统的复杂性是如何进化出来的？这个进化生物学中最重要的问题目前仍未得到解决。在这一章我们看到，对生物复杂度的完全理解现在才刚刚开始。我们也看到，对复杂性的进化的理解已经迈出了一些重要的步伐。其中包括一些人提出的“扩展综合（extended Synthesis）”，这个理论认为自然选择仍然很重要，但是其他因素——

历史偶然、发育制约（developmental constraint）和自组织——也和自然选择一样重要。进化论者由于受到宗教极端主义的攻击，尤其是在美国，并且经常处于守势，因而不愿意承认自然选择可能不是故事的全部。对于这种窘境，生物学家霍泽尔（Guy Hoelzer）、佩珀（John Pepper）和埃里克·斯密斯（Eric Smith）评论道："加入捍卫达尔文主义的战斗现在已经成为进化生物学家必需的社会责任，[312] 但这种文化范式也带来了对科学的负面影响。对进化过程进行新的解释，以对自然选择进行补充，也会激起同样的不经思索的反对立场。"

进化生物学家麦舒（Dan Mc Shea）教了我一个思考这些问题的好办法。[313] 他将进化论者分为三类：适应主义者，认为自然选择才是主要的；历史主义者，相信历史偶然导致了许多进化变化；以及考夫曼这样的结构主义者，关注的是组织结构如何能没有自然选择也能产生。只有这三类人证明他们的研究能够成为一个整体时，进化论才能统一。

麦舒也为我进行了乐观的展望："进化生物学处于知识混沌状态中，[314] 但这个知识混沌极富成效。"

5 尾声

我要将混沌放入十四行诗，[315]

让他留在那里，让他从此逃遁。

如果他幸运，让他变形、伪装，

洪水、火焰、恶魔——他机敏的谋划，

在甜美秩序的严格界限中游刃有余，

在道貌岸然的强暴中，

我抓住了他的本质和易变的模样，

直到他与秩序混合交叠。

日复一日，年复一年，

他傲慢自大，我们卑躬屈膝：

我抓住了他。他只不过是

简单却没有被理解之物。

我甚至不会逼他忏悔；

或是坦白。我只是将他驯服。

——埃德娜·圣文森特·米莱（Edna St.Vincent Millay），

《采获丰收：新诗集》（Mine the Harvest: A Collection of New Poems）

第19章　复杂性科学的过去和未来

1995年，科学记者约翰·霍根（John　Horgan）在世界顶级科普杂志《科学美国人》上发表了一篇文章[316]，攻击整个复杂系统研究，尤其是圣塔菲研究所。他的文章标题“复杂性是不是骗局？”被印在杂志封面上（图19.1）。

这篇文章主要有两个看法。首先，在霍根看来，复杂系统研究不太可能发现什么有用的一般性原则；其次，他相信计算机建模的盛行使得复杂性成了“与事实无关的科学”。另外，文中还有几处过激攻击，将复杂性称为“流行科学”，称其研究者为“复杂学家（complexologist）”。霍根揣测“复杂性”这个词没有什么意义，我们留着它只是因为它具有“公关价值”。

更过分的是，霍根在文中引用我的话说：“在某种程度上你可以说所有复杂系统都体现了同样的根本原则，但我不认为那会很有用。”我真的这样说过吗？我怀疑。这样说的背景是什么？我是表示认同这种观点吗？霍根通过电话采访了我一个多钟头，我说了很多；他单单选了最负面的评论。当时我还没有多少同科学记者打交道的经验，我感到非常生气。

SCIENTIFIC AMERICAN

JUNE 1995
$3.95

The world's strongest magnets.

Is complexity a sham?

Found: 2,000-year-old blueprint.



***Picky wildflowers** choose which pollen to accept.*

图19.1　《科学美国人》杂志在封面上对复杂性进行“羞辱”（Rosemary Volpe封面设计，经许可重印）

我给《科学美国人》的编辑写了一封信表达我的愤怒，列出了我认为霍根的文章中所有错误和不公平的地方。当然，我的许多同事也这样做了；杂志只发表了其中一封信，不是我那封。

这件事给了我一些教训：和记者打交道要非常小心。但也促使我更努力细致地思考“一般性原则”的概念，以及这个概念可能的意义。霍根将他的文章扩充成了一本同样充斥着敌意的书，《科学的终结》[317]，他在书中提出所有真正重要的科学发现都已经完成了，以后不会再有了。他在《科学美国人》上写的复杂性文章扩充成了书的一章，并且包括以下消极论断：“混沌、复杂性和人工生命的研究还将继续……但他们不会得到任何对自然的伟大洞察——更不可能与达尔文理论或量子力学相提并论。”

霍根有可能是对的吗？发现所有复杂系统的一般性原理或“统一理论”的目标会是徒劳吗？

统一理论和一般性原理

统一理论[unified theory，或大统一理论（Grand Unified Theory），缩写为GUT]通常指物理学的一个目标：用一个理论统一宇宙中的基本力。弦论就是对GUT的尝试，但对于弦论是不是正确的，甚至GUT是不是存在，在物理学界还没有达成共识。

假如弦论最后被发现是对的——它是物理学长期追寻的GUT，那将会是极其重大的成就，但那不会是科学的终结，更不会是复杂系统科学的终结。让我们感兴趣的复杂系统行为是无法在基本粒子或是十维的弦这样的层面上进行理解的。即使现实世界中的一切都是由基本粒子组成，它们也不是解释复杂性的适当词汇。这就好像在问到“为什么逻辑斯蒂映射是混沌的”时回答说：“因为$x_{t+1}=Rx_t(1-x_t)$。”在某种意义上，对混沌的解释确实包含在这个等式中，就好像免疫系统的行为在某种意义上包含在物理的大统一理论中一样。但这并不构成人类的理解，而这才是科学的终极目标。物理学家克鲁奇菲尔德、法墨尔、帕卡德和罗伯特·肖（Robert Shaw）对此说得非常好：“希望物理学能彻底理解基

本力从而完结的想法是没有根基的。[318] 一个尺度上组分的相互作用会导致更大尺度上复杂的全局行为，而这种行为一般无法从个体组成的知识中演绎出来。”爱因斯坦也曾开玩笑说：“人们相爱不能怪万有引力。”[319]

如果复杂系统的统一理论不是基础物理学，那会是什么呢？又有没有呢？大部分复杂系统研究者可能都会说寻求复杂性的统一理论现在还为时尚早。物理科学发展了两千多年，在认识到质量和能量这两种主要“要素”后很久，才由爱因斯坦用$E=mc^2$ 将它们统一。物理学还认识到自然界的四种基本力，并且统一了其中至少三种。质量、能量和力，以及它们背后的基本粒子，是物理学理论的基本要素。

对于复杂系统，我们甚至不知道它的基本“要素”或基本“力”是什么；除非你已经知道了统一理论的概念组成或基本要素，否则谈论统一理论没有什么意义。

生态和昆虫学家戈登则表达了以下观点：

> 继控制隐喻（metaphors of control）之后 [320] ，关于复杂性、自组织和涌现的思想——整体大于部分之和——也开始流行起来。但这些解释都只是障眼法，仅仅给出了一些我们无法解释的名词；它们给我的感觉就好比物理学家用等式中两项相等解释粒子的行为，无法让人满意。也许存在复杂系统的一般性理论，但是很明显目前还没有。关注具体系统的细节是理解自组织系统动力学更好的途径。这样可以发现是否存在一般性规律……希望用一般性原理来解释自然界中发现的各式各样复杂系统的规律，这会让我们忽视与模型不符的现象。多了解这类系统的具体特性，就能发现在各系统之间哪些类推有效，哪些类推又无效。

确实有很多一般性原理不是很有用，例如，“所有复杂系统都具有涌现性质（emergent properties）”，因为就像戈登说的，它们给出的是“我们无法解释的名词”。我想，这就是我在说霍根引用的话时所表示的意思。没有单一的原理可以适用于所有复杂系统，我认为从这个意义上说戈登是对的。

用共性而不是一般性原理来称呼可能更好：它们为一些系统或现象

的机制给出了新的理解或概念，如果没有它们，很难通过分别研究这些系统或现象然后进行类比来厘清。

共性的发现可以是复杂性研究反馈环的一部分：由具体复杂系统的知识总结出共性，反过来又为理解具体系统提供了思想。具体的细节与共性相互启发、约束和丰富对方。

听起来很不错，但有没有例子呢？各种文献中提出了很多共性或普适原理，这本书中我们也看到了一些：混沌系统的普适性质；冯·诺依曼的自复制原理；霍兰德的搜索与开发平衡原理；阿克塞尔罗德的合作进化的基本条件；沃尔夫勒姆的计算等价性原理；巴拉巴西和艾伯特提出偏好附连是真实网络发展的普遍机制；韦斯特、布朗和恩奎斯特提出用分形循环网络解释比例关系；等等。还有很多，受篇幅限制我们无法一一列举。

在第12章我壮着胆子提出了分散系统中适应性信息处理的一些共通原理。我不知道戈登是否会同意，但我相信对于人们研究与我提到的系统类似的系统应该会有用——这些原理也许能给他们一些新的思想，帮助他们理解所研究的系统。例如，我提出了“随机性和或然性很关键”。最近我在一次演讲中阐述了这些原理，听众中一位神经学家就跟着推测了大脑中随机性的可能来源，以及起到了什么作用。听演讲的人中一些人从没有从这个角度思考过大脑，因此这个思想稍微改变了他们的观念，也许在他们以后的研究中就会用到这些新概念。

另一方面，从具体系统到共性也会有反馈。也是在那次演讲中，一些人举了一些复杂适应系统的例子，他们认为这些系统并不符合我的全部原理，这驱使我重新思考我的观点的通用性。正如戈登指出的，我们应当注意不要忽视“与模型不符的现象”。当然，对自然现象的认识有时候也会有错误，而共性也许能引导我们进行辨析。据说爱因斯坦——他是杰出的理论大师——曾说过：“如果事实与理论不符，就改变事实。”当然，这取决于是什么理论和事实。理论越是稳固，你就越应当怀疑与之相抵触的事实，反过来如果与之相抵触的事实越是有根据，你就越应当怀疑你提出的理论。这就是科学的本性——永无止境的提议和质疑。

复杂系统研究的根源

对复杂系统共性的寻找有很长的历史，特别是在物理学中，但发展最快的阶段还是在计算机发明以后。20世纪40年代初，一些科学家提出计算机与动物之间有很强的相似性。

20世纪40年代，以赛亚·梅西基金会（Josiah Macy，Jr.Foundation）资助了一系列交叉科学会议，主题很有趣，包括“生物和社会系统中的反馈机制和循环因果系统”“社会的目的论机理”以及“目的论机理与循环因果系统”。这些会议是由一小群探寻各种复杂系统共性的科学家和数学家组织的。这个团体的主要推动者是数学家维纳（图19.2），他在第二次世界大战期间研究高射炮的控制，这段研究经历让他认识到，无论是生物还是工程中的复杂系统，研究的关键都不再是质量、能量和力这些物理学概念，而是反馈、控制、信息、通信和目的（或“目的性”）等概念。

梅西基金会系列会议聚集了当时许多杰出人物，除了维纳，还有冯·诺依曼、麦卡洛克（Warren Mc Culloch）、米德（Margaret Mead）、贝特森、香农、阿什比（W.Ross Ashby）等人。这些会议促使维纳提出了一门新的学科，控制论[321]（cybernetics），这个词来自希腊语的“舵手”一词，也就是船的操控者。维纳将控制论归结为“整个控制和通信的理论[322]，无论是关于机械还是动物”。

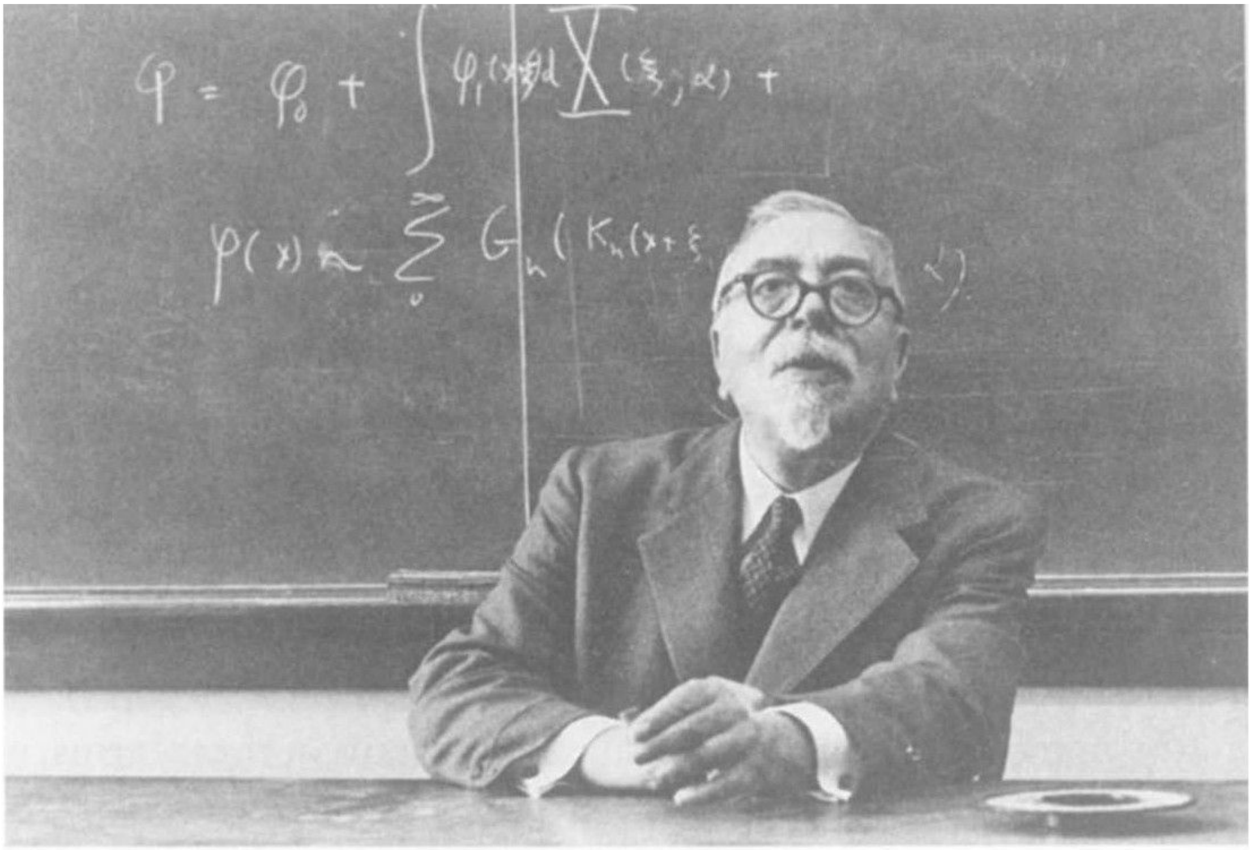


图19.2　罗伯特·维纳（1894—1964）（AIP Emilio Segre Visual Archives）

这个松散的控制论团体关注的主题在这本书中都出现过。他们想知道：信息和计算是什么？它们在生物中是如何表现的？生物与机器有什么相似之处？反馈在复杂行为中起什么作用？信息处理是如何产生出意义和目的的？

控制论团体在生物与机器的类似性上做了许多重要工作。例如冯·诺依曼的自复制自动机将信息与繁殖联系到了一起；阿什比的《大脑设计》[323]提出将动力学、信息和反馈应用到神经科学和心理学。麦卡洛克和皮茨（Walter Pitts）提出了神经元模型作为逻辑器件，[324]引发了后来神经网络的研究；米德和贝特森将控制论的思想应用到心理学和人类学；[325]维纳的著作《控制论》和《人有人的用处》[326]则试图为这个领域和许多相关学科提供统一的认识。这些成就只是其中部分例子，它们的影响延续至今。

控制论的研究在当时既有人拥护也有人反对。拥护者认为其开创了科学的新时代；批评者则认为它没有什么用，因为太过宽泛模糊，缺乏

严格的理论基础。人类学家贝特森认同前一观点，他写道："我一生经历的最重要的两次历史事件是凡尔赛条约的签订和控制论的发现。" [327] 而生物学家和诺贝尔奖获得者德尔布吕克（Max Delbrück）则认为他参加的控制论会议"极为空洞无物" [328] 。决策论学家萨维奇（Leonard Savage）说得客气一点，他说后期的一次梅西基金会会议是"非常精英的团体在一起闲谈" [329] 。

控制论主义者参加会议的热情逐渐消退，这个领域本身却繁荣起来。科学史学家艾斯普瑞（William Aspray）研究了控制论运动，他写道："最后维纳统一控制和通信科学的愿望没有实现 [330] 。就像其中一位参与者评论的，控制论'宽泛而缺乏实质'。涵盖的主题过于松散，理论工具又过于贫乏笨拙，无法实现维纳所期望的统一。"

还有一个寻找共性的类似尝试，就是所谓的一般系统论 [331] ，20世纪50年代由贝塔朗菲（Ludwig Von Bertalanffy）发起，他将其描述为"对一般性'系统'有效的原则进行形式化和演绎" [332] 。系统是在非常一般性的意义上进行定义：由相互作用的组分组成的集合，组分通过相互作用一起产生出某种形式的系统及行为。当然，这什么都可以描述。一般系统论者最感兴趣的是生物系统的一般性质。系统论学家拉普波特将一般系统论（应用到生物系统、社会系统和其他复杂系统）的主线描述为在变化中保持的一致性，有组织的复杂性以及目标导向性。生物学家马图拉纳（Humberto Maturana）和维埃拉（Francisco Varela）试图用自创生（autopoiesis，或"自我建构"）的概念统一前两条主线， [333] 这个概念表示自我维持的过程，系统（例如一个生物细胞）作为一个整体运转，不断产生出系统本身的构成组分（例如细胞的部件）。对于马图拉纳、维埃拉和他们的许多追随者来说，自创生即便不是唯一，也是一个重要的生命特性。

同控制论的研究目标一样，这些思想非常吸引人，但是建构严格的数学框架来解释和预测这类系统重要共性的尝试没有获得普遍成功。然而在这些尝试中提出的核心科学问题形成了一些现代科学和工程领域的基础。人工智能、人工生命、系统生态学、系统生物学、神经网络、系统分析、控制理论和复杂性科学都是由这些控制论学家和一般系统论学者播下的种子发展而来。对控制论和一般系统论的研究仍然很活跃，但基本已经被这些从中衍生出来的学科掩盖了。

后来的一些针对复杂系统一般性理论的尝试来自物理学。例如，哈肯（Hermann Haken）的协同学和普里高津（Ilya Prigogine）的耗散结构和非平衡系统理论，[334] 都是试图结合热力学、动力系统理论和“临界现象”理论来解释湍流、复杂化学反应这类物理系统以及生物系统的自组织。特别是，普里高津的目标是确定“复杂性的词汇表”：用普里高津和他的同事尼古拉斯（Grégoire Nicolis）的话说，“涉及在各种现象中反复遇到的机制的一系列概念[335]；包括非平衡性、稳定性、分岔和对称破缺，以及长程有序（long-range order）……我们相信这些是一个新的科学词汇表的基本组成”。研究在不断沿着这些方向进行，但直到目前仍然没有产生出普里高津所预想的那种具有一致性和一般性的复杂性词汇表，更不要说能将这些不同的概念统一到一起，解释自然界中的复杂性的一般性理论。

五个问题

从这本书所涵盖的主题之广泛可以看到，现代复杂系统科学仍然没有统一成一个整体，而是松散的大杂烩，其中有一些相互重叠的概念。目前在这个标题下统一的只有共同的问题和方法，以及超越早期研究中不那么严格的类比特性，得到更严格的数学和实验的渴望。对于现代复杂系统科学相对于以前的尝试有何新的贡献，或者有没有贡献，存在很多争议。它有多成功呢？

对这个问题有各种看法。最近，一位名叫吉尔森逊（Carlos Gershenson）的学者向他的一些同行（其中包括我）分发了一份复杂系统问题表，并计划在名为《复杂性：5个问题》（Complexity：5 Questions）的书中发表这些回应。问题如下：

1.你为何会研究复杂系统？

2.你怎样定义复杂性？

3.你喜欢的复杂性方面/概念是什么？

4.在你看来，复杂性最成问题的方面/概念是什么？

5.你如何看待复杂性的未来？

目前我看到了其中14份回应。虽然表达的观点多种多样，但还是涌现出了一些共同的想法。大部分人认为复杂性的“普适定律”的可能性过于野心勃勃或过于模糊不清。而且，大部分人都认为定义复杂性是这个领域最成问题的方面，可能根本就是错误的目标。许多人认为复杂性一词没有意义；一些人甚至避免使用它。大部分人不认为已经存在“复杂性科学”，至少不是在科学一词的通常意义上——复杂系统似乎是一个四分五裂的学科，而不是统一的整体。

最后，有少数人担心复杂系统领域会遭遇与控制论等相关尝试同样的命运——也就是说，它将阐明不同系统之间有趣的类似之处，而不会得出一致而严格的数学理论，从而解释和预测它们的行为。

不过，虽然对当前的复杂系统研究有这些消极看法，大部分人对于这个领域以及其对科学已经产生和将要产生的贡献还是抱以高度热情。在生命科学、大脑科学和社会科学中，科学家们研究得越深入，发现的复杂现象就越多。新的技术手段使得这样的发现越来越多，这些发现极需有新的概念和理论来解释复杂性的来源和机制。这些发现需要科学做出改变，抓住复杂系统研究中出现的问题。事实上，在本书前面的例子中可以看到，近年来复杂性科学的主题和结果已经触及几乎所有科学领域，而且像生物学和社会学这样的研究领域已经被这些思想深深改变了。不仅如此，一位学者这样说道：“我认为复杂性科学的一些形式正在改变整个科学思想。”一些参与调查的人也表达了类似想法。

除了布朗、恩奎斯特和韦斯特的代谢比例研究和阿克塞尔罗德等重要的具体发现，到目前为止复杂系统研究最有意义的贡献也许是对许多长期持有的科学假设提出了质疑，并且发展出了将复杂问题概念化的新方法。混沌告诉了我们看上去行为随机的系统并不一定是因为有内在的随机性；遗传学的新发现对基因变化在进化中的作用形成了挑战；对随机和自组织的作用的新认识挑战了将自然选择作为进化的核心力量的观念。非线性、分散控制、网络、层次、分布式反馈、信息的统计表示、本质的随机性，这些思想的重要性在科学界和大众中都逐渐被认识到。

新的概念体系经常需要对存在的概念进行拓宽。这本书中我们看到了信息和计算的概念如何被拓展到涵盖生命系统，甚至复杂社会系统；适应和进化的概念如何被拓展到生物王国之外；生命和智能的观念如何被拓展到自复制机器和进行类比的计算机程序。

这种思考方式逐渐进入主流科学。我在SFI暑期学校与年轻的研究生和博士后进行交流时清楚地看到这一点。20世纪90年代初，同学们对暑期学校讲授的新思想和新颖的科学世界观都极为兴奋。但进入21世纪以后，主要得益于SFI等研究机构的大力宣传，这些思想和世界观已经渗入许多学科的文化，逐渐习以为常了，有时候甚至对复杂系统变得如此“主流”感到失望。我想，这应当视为一种成功。最后，复杂系统研究强调多学科合作，现在看来这对目前那些最为重要的科学问题的研究非常关键。

复杂性的未来，等待卡诺

在我看来，复杂系统科学正分化成两个独立的方向。沿其中一个方向，复杂性研究的思想和工具被提炼出来，并应用到更广泛的领域。在这本书中我们已经看到，相似的思想和工具被应用到物理学、生物学、流行病学、社会学、政治学和计算机科学等截然不同的领域。在一些我没有讨论的领域，如神经科学、经济学、生态学、气候学和医学，复杂系统的思想也占据了越来越重要的地位——复杂性和交叉科学的种子撒播得越来越远。

另一个方向则更具争议，它从更高的层面上来审视这些领域，寻求解释性和预测性的数学理论，将复杂系统之间的共性严格化，并且能解释和预测涌现现象。

在我参加的一次复杂性会议上，对于这个领域应当向哪个方向发展，进行了一次热烈的讨论。在一时的失落气氛中，一位与会者说：“‘复杂性’曾让人兴奋过，但现在已经死了。我们应当另起炉灶。”

我们应当如何描述它呢？现在也许清楚了，这才是问题的关键——我们没有合适的词汇表来精确描述我们所研究的对象。我们用复杂性、自组织和涌现来描述我们感兴趣的系统的共同现象，但是我们还是不能以更严格的方法刻画这些共性。我们需要新的词汇表，不仅能抓住自组织和涌现的概念构成，还能解释它们如何涵盖所谓的功能性、目的或意义（参见第12章）。这些不清晰的词汇需要用新的更清晰的词汇来定义，以反映出对所研究的现象的新理解。就像我在书中介绍的，复杂系统的许多研究都涉及对来自动力学、信息、计算和进化的概念进行整合。应当通过这种整合形成新的概念词汇表和新的数学。数学家斯托加

茨这样说道："我认为我们可能缺乏与微积分相当的新概念体系[336]，能根据复杂系统的无数相互作用得到其结果的方法。这种超级微积分，即使告诉了我们，也有可能超出人类的理解能力。到底怎样我们不得而知。"

要想理解、预测或是引导和控制具有涌现性质的自组织系统，就必须有适当的概念词汇表和适当的数学。发展出这样的概念和数学工具在过去和现在都是复杂系统科学所面临的最大挑战。

这个领域有个笑话，说我们是在"等待卡诺"。卡诺（Sadi Carnot，图19.3）是19世纪初的一位物理学家，他提出了热力学的一些关键概念。与之类似，我们也在等待出现适当的概念和数学来描述我们在自然界看到的各种形式的复杂性。

要实现这个目标我们更需要一位牛顿式的人物。我们现在所面临的概念问题，就类似于微积分发明之前牛顿所面临的问题。在牛顿的传记中，科学作家格雷克（James Gleick）这样描述："他受困于语言的混乱[337]——有些词汇定义不清，有些词汇甚至还没有出现……牛顿相信，只要他能找到合适的词汇，他就能引领整个运动科学。……"通过发明微积分，牛顿最终创造了所需的词汇。借助于无穷小、微分、积分和极限等概念，微积分为严格描述变化和运动提供了数学语言。这些概念在数学中已经存在，但是不完整；牛顿能够发现它们之间的关联，并且构建出和谐统一的宏大建筑将它们结合到一起，让它们彻底一般化。这幢宏大的建筑使得牛顿能够创造出动力学体系。



图19.3　卡诺（1796—1832）[布瓦利（Boilly）版画，摄影学会，柏林，由美国物理学会西格尔图像档案提供，哈佛大学藏品。]

我们能够类似地发明出复杂性的微积分吗——一种能抓住复杂系统的自组织、涌现行为和适应性的起源和机制的数学语言？一些人已经开始着手于这项宏伟计划。例如，第10章曾介绍过，沃尔夫勒姆正在以元胞自动机中的动力学和计算为基础，创造他所认为的新的基础性的自然理论。前面提到，普里高津和他的追随者曾尝试用一些物理学概念作为基础建立复杂性理论。物理学家巴克（Per Bak）在动力系统理论和相变概念的基础上提出了自组织临界性的概念，[338] 并将其作为自组织和涌现的一般性理论。物理学家克鲁奇菲尔德提出了计算力学（computational mechanics），[339] 将动力系统、计算理论和统计推断理论结合到一起，解释复杂和适应性行为的涌现和结构。

不过这些方法，以及我没有提到的其他一些方法，都还远没有成为被广泛接受的复杂系统的解释性理论。它们都包含有一些重要的新思想，目前仍然是活跃的研究领域。当然，目前仍然不清楚是否存在这样一个理论；有可能不同系统中的复杂性的产生和运作过程完全不同。在这本书中我介绍了一些复杂系统理论的可能片段，分别涉及信息、计算、动力学和进化等领域。需要做的是发现它们的内在关联，并将它们融合成协调一致的整体——也许可以称其为“复杂性背后的简单性”[340]。

虽然这本书中介绍的许多科学仍然处于初期阶段，但对我来说，实现这种远大目标的前景正是复杂系统研究真正的迷人之处。有一件事情很清楚：追寻这些目标，要具有在知识上冒险和不惧失败的精神，敢于超越主流科学，进入疑点重重的未知领域，伟大的科学都是这样的。借用作家和探险家纪德（André Gide）的一句话：“不敢远离海岸线，就别想发现新大陆[341]。”朋友们，让我们一起向复杂性的新疆域进发吧。

附录　访谈——梅拉妮·米歇尔谈复杂性

《泛在》杂志（*Ubiquity*）2011年4月

梅拉尼·米歇尔，1990年在密歇根大学获博士学位，导师是侯世达。她曾在圣塔菲研究所和俄勒冈研究院任职，2004年加入波特兰州立大学，现为波特兰州立大学计算机科学教授和圣塔菲研究所外聘教授，著有《复杂》一书。这本书由牛津大学出版社出版，内容引人入胜、富有启发性，被评为亚马逊网站2009年度十佳科学图书。米歇尔的研究范围涵盖人工智能、机器学习、生物启发计算、认知科学和复杂系统。

《泛在》：先问一个简单的问题，什么是“复杂性”？

米歇尔：这个问题“看似简单”——其实是最复杂的问题！复杂性研究之所以产生，是因为一些学者强烈地感觉到，一些高度“复杂”的自然、社会和技术系统之间具有深刻的相似性。这种系统的例子包括大脑、免疫系统、细胞、昆虫社会、经济、万维网，等等。说它们“相似”，并不是说必然存在掌控这些不同系统的唯一的一组原理，而是说所有这些系统都表现出“适应性的”、“类似生命的”、“智能性的”和“涌现性的”行为。这些术语都没有精确的定义，也使得目前还不可能形式化地定义“复杂系统”。

有一个通俗的复习系统定义：由大量相互作用的组分组成的系统，与整个系统比起来，组分相对简单，没有中央控制，组分之间也没有全局性的通信，并且组分的相互作用导致了复杂行为。这里“复杂行为”指的是前面列出的那些术语（适应性、涌现，等等）。

《泛在》：存在复杂性科学吗？

米歇尔：我认为复杂性研究是不同学科的松散组合，研究复杂系统并寻求厘清这些系统之间的共同原则。100多年前，哲学家和心理学家威廉·詹姆士曾说过，心理学还不是科学，只是“有希望成为科学”。我认为这对于今天的复杂性研究来说也同样成立。我个人尽量避免使用“复杂性科学”（complexity science）一词，而是用“复杂性研究”（the

sciences of complexity）。

《泛在》：圣塔菲研究所（SFI）于1984年成立，是复杂系统研究的中心。你是如何加入SFI的？

米歇尔：我当时是密歇根大学的研究生，攻读计算机科学的博士学位，导师是侯世达。我选了霍兰德教授的“遗传算法”课，他是SFI的早期成员之一。我对这个领域产生了浓厚的兴趣，霍兰德邀请我到SFI访问了一个暑假。我迷上了这里，很想找机会再去。在暑期访问大约一年后，我获得了博士学位。当时SFI正好有一个主持“适应性计算”项目的职位。霍兰德再次推荐了我，我在那里全职工作了一段时间，并最终成为研究所的一员。

《泛在》：你又是怎么去波特兰州立大学的？

米歇尔：我在圣塔菲的时候，研究所规定任职期限为五年，不可延期。期限结束后，我和我丈夫想搬到波特兰，我就申请了波特兰的俄勒冈研究院计算机科学的职位。俄勒冈研究院最近与俄勒冈卫生科技大学合并了。我在俄勒冈研究院的第二年，由于研究院目标的一系列变化，我们系的大部分教员，包括我自己，都受邀加入了波特兰州立大学（PSU）的计算机科学系。这是两个系很有趣而且史无前例的合并，PSU的计算机科学系扩大了一倍。

《泛在》：我们马上还会谈到计算机科学，不过能不能首先请你总结一下对复杂性研究有贡献的其他学科？

米歇尔：有很多学科都有贡献，而且数量还在不断增长！可以说几乎科学的所有主要分支对复杂系统研究都有一定的贡献，社会科学、历史和哲学的各分支也是一样。我不能确定边界在哪里。

《泛在》：你的书详细讨论了生物学与复杂性的互动，并将生物学与计算联系起来。信息处理在生命系统中扮演了什么角色？

米歇尔：对这个问题的回答可以写一本书（而且我确实打算以后写一本！）。简单地说就是，信息处理是描述生命系统行为的另一种方式；也就是说，不同于典型的生物学文献中的那种方式。信息处理或计算的框架，能帮助我们统一在生命系统中发现的一些不同特性。在《复

杂》中，我讨论了蚁群、免疫系统和细胞代谢，通过将它们的行为都视为“计算”来描述它们的相似性，这样我就能提出在这些非常不一样的生物系统中一些共同的信息处理原理。但是有必要指出，生命系统中的“信息处理”和“计算”的完整概念仍然相当模糊——许多人用这些术语来描述生物现象，但对其的定义或形式化并没有达成共识。有时候很难清楚地知道人们谈论的是什么。

《泛在》：计算机科学应该为这个一般性领域做些什么？

米歇尔：计算机科学可以有许多贡献。计算机仿真是研究复杂系统的核心方法。高性能计算机建模工具和方法对于这个领域的推进具有绝对的重要性。但我认为，计算机科学还有更重要的作用，那就是为思考自然界中的信息处理提供形式化框架。复杂和适应系统的一个标志是以复杂的方式处理信息的能力。例如，整个生物学领域都越来越多地用信息处理的观念来作为理解适应性行为的框架。

我相信计算机科学以及更一般性的计算理论，最终会将目前还不正式而且模糊的信息处理概念形式化。我个人认为信息处理将会成为理解生命系统的一个统一框架。

《泛在》：你的博士论文中介绍了你写的一个能进行类比的计算机程序。写得相当漂亮！

米歇尔：谢谢。这是我仍然在研究的课题。

《泛在》：进行类比与复杂性有什么关系？

米歇尔：进行类比是认知的核心。在这里无法深入阐释，我写的《作为感知的类比》（*Analogy-Making as Perception*），侯世达的《流动的概念与创造性类比》（*Fluid Concepts and Creative Analogies*）和即将出版的《表面与本质》（*Surfaces and Essences*，与伊曼纽尔·桑德合著），都是关于这个主题，书中描述了类比存在于智能核心的许多方式。侯世达和我建立的人们如何进行类比的模型中体现了它与复杂性的关系。这个模型涉及蚁群、免疫系统等复杂生命系统所共有的一些适应性信息处理的重要机制。在我的新书中用了一章来描述其中的关联。

《泛在》：网络科学有什么用途？

米歇尔：网络科学试图研究各学科中的各种网络，并给出共通的原理和方法。当然，网络以各种面目被研究了数百年：图论是对网络的数学研究；社会学家和社会心理学家研究社会网络；工程师研究电力网络和互联网这类技术网络；生物学家研究食物网；遗传学家研究基因调控网络；等等。但大多数时候各学科之间没有相互沟通。最近人们才开始发现这些不同系统之间有趣的共性。例如，20世纪90年代末，瓦茨和斯托加茨给出了网络的“小世界性”的数学定义，并证明了电力网络、线虫的神经网络和演员的社会网络都具有这种特性。此后小世界性得到了深入研究，并在许多学科研究的网络中都有发现。

同样也是在20世纪90年代末，巴拉巴西和艾伯特提出了“无尺度网络”的概念（本质上是具有分形结构的网络），并证明了许多自然和技术网络都具有无尺度结构。由于许多（也许是所有）复杂系统都可以被视为网络，个体（节点）与有限数量的其他个体进行通信（连接），因此网络的交叉学科研究有可能揭示复杂网络的普遍共性。

《复杂》中用了几章来探讨网络科学的内容和影响。巴拉巴西的《链接》（*Linked* ）[342] 和瓦茨的《六度》（*Six Degrees* ）都对这门新涌现的学科有很好的介绍。更专业的介绍可以参见纽曼最近写的《网络引论》（*Networks*：*An Introduction* ）一书。

《泛在》：有一个问题是“无尺度”是否适用于互联网。大卫·奥尔德逊在2009年的《泛在》访谈中提出，互联网以及谷歌云这类互联网上的子系统具有针对性的工程设计，比无尺度模型所预测的更具有稳健性。

米歇尔：奥尔德逊的例子（路由跟踪数据、谷歌服务器）指的是互联网（由服务器和其间的通信连接组成），与我前面提出的万维网的例子不是一个概念（万维网是由超链接组成的逻辑结构），而万维网才被许多人认为具有无尺度性。但大的问题仍然成立；总体上很难判断一个大型网络是否确实是“无尺度的”，还是具有某种其他结构。精确地说，“无尺度”一词指的是具有连接度幂律分布的数学性质。在现实世界中，这个性质只能被近似满足，没有绝对的“无尺度”网络，就像自然界中没有完美的分形一样。因此问题是：在何种程度上我们可以说网络是无尺度的，这种近似对于理解网络的行为有用吗？这是网络科学文献中许多争议的主题，但是有许多经验研究表明，包括万维网在内的许多自

然和技术网络具有（近似的）无尺度特性。

《泛在》：依你看，复杂性研究以及计算机科学的总体目标应当是什么？

米歇尔：我认为有两个关联的目标，都还远远没有达到。

首先是发现不同的复杂系统之间的共同原理，从而得到对这些系统的洞察，并产生出分析这些系统的新方法。无尺度网络就是共同原理的一个例子，我们刚刚讨论过。生物学家则借鉴了谷歌的网页排名算法（一种利用了万维网的无尺度结构的计算方法），用来研究食物网中不同物种的重要性，从而更好地认识灭绝的风险，这就是新的分析方法的一个例子。（相关研究可以参见一篇文章：Allesina S，*Pascual M, Googling Food Webs*：*Can an Eigenvector Measure Species'Importance for Coextinctions?* PLo S Comput Biol，5（9），2009.）

其次，更具雄心的目标也许是发展出数学理论，以一般性的方式描述复杂性，并对许多不同系统的现象进行解释和预测。例如，有了这样的理论，就有可能以形式化的方式明确昆虫群体、经济系统和大脑等复杂系统所共有的动力学、适应、集体决策和控制以及“智能”背后的机制。这样一个理论应当结合动力系统理论、计算理论、统计物理、随机过程、控制理论、决策论等领域的理论研究。对于是否存在这样一个理论目前都还不清楚，更不要说这个理论是什么样子。

数学家斯托加茨称这个目标为寻找“复杂性的微积分”。从一些方面来看，这个类比很贴切：牛顿、莱布尼茨等人寻找的就是运动的一般性理论，以解释和预测服从物理力的任意物体的动力学，无论是地上的还是天上的。在牛顿之前的时代，这个理论已经有了一些片段（例如，已经存在“无穷小”“导数”“积分”等概念），但还没有人将这些片段合到一起，给出完整的一般性理论来解释以前没有统一认识的各种现象。与此类似，我们也有与复杂系统有关的各种理论片段，但还没有人知道如何将它们合到一起，产生出某种更具一般性和统一性的理论。

《泛在》：你真的认为复杂性的微积分会出现吗？

米歇尔：这种一般性理论曾是20世纪40年代和50年代控制论运动的圣杯；从许多方面来说，目前的复杂系统研究是那场运动的延续。一些

批评意见认为复杂系统会与控制论有类似的命运：不能超越建设性的隐喻和分散的片段，无法给出更严格和有用的框架。我个人的看法要乐观一些，但也不好说……复杂性研究还很年轻，还具有很多革新性发展的潜力，我希望它能不断吸引世界上一些最具创造性的年轻学家加入。

http：//ubiquity. acm.org/article.cfm?id=1967047

感谢ACM和《泛在》提供中文版权

Copyright 2011 ACM, Inc.

参考文献

Achacoso, T. B.and Yamamoto, W.S.*AY's Neuroanatomyof C.Elegansfor Computation.* Boca Raton, FL：CRC Press，1991.

Adami, C.*Introduction to Artifcial Life*，Springer，1998.

Agutter P. S and Wheatley D.N.Metabolic scaling：Consensus or Controversy?*Theoretical Biology and Medical Modeling*，18，2004，pp.283—289.

Albert, R. and Barabási, A-L.Statistical mechanics of complex networks.*Reviewsof Modern Physics*，74，pp.48—97，2002.

Ashby, H. R.*Design for a Brain.* New York：Wiley，1954.

Aspray, W.*John von Neumann and the Origins of Modern Computing.* Cambridge, MA：MIT Press，1990.

Aubin, D. and Dalmedico, A.D.Writing the history of dynamical systems and chaos：*Longue Durée* and revolution, disciplines, and cultures.*Historia Mathematica*，29，2002，pp.273—339.

Axelrod, R.*The Evolution of Cooperation.* New York：Basic Books，1984.

Axelrod, R. An evolutionary approach to norms.*American Political Science Review*，80（4），1986，pp.1095—1111.

Axelrod R. Advancing the art of simulation in the social sciences.In Conte, R.，Hegselmann, R.，Terna, P.（editors）.*Simulating Social Phenomena.*（Lecture Notes in Economics and Mathematical Systems 456）.Berlin：Springer-Verlag，1997.

Bak, P.*How Nature Works*：*The Scienceof Self-Organized Criticality*. New York：Springer，1996.

Barabási, A.-L.*Linked*：*The New Science of Networks.* Cambridge, MA：Perseus，2002.

Barabási, A.-L.and Albert, R.Emergence of scaling in random networks.*Science*，286，1999，pp.509—512.

Barabási, A.-L.and Oltvai, Z.N.Network biology：Understanding the cell’s functional organization.*Nature Reviews*：*Genetics*，5，2004，pp.101—113.

Barrett, P.（editor）.*Charles Darwin's Notebooks*，1836—1844：*Geology, Transmutationof Species, Metaphysical Enquiries.* Ithaca, NY：Cornell University Press，1987

Bassett, D. S.and Bullmore, D.Small-world brain networks.*The Neuroscientist*，12，2006，pp.512—523.

Bateson, G. H.*Mind and Nature*：*A Necessary Unity*. Skokie, IL：Hampton Press，1979.

Beinhocker, E. D.The*Origin of Wealth*：*Evolution, Complexity, and the Radical Remaking of Economics.* Cambridge, MA：Harvard Business School Press，2006.

Bennett, C. H.The thermodynamics of computation-a review.*International Joural of Theoretical Physics*，21，1982，PP.905—940.

Bennett, C. H.Dissipation, information, computational complexity and the definition of organization.In Pines, D.（editor），*Emerging Syntheses in Science.* Redwood City, CA：Addison-Wesley，1985，pp.215—233.

Bennett, C. H.How to define complexity in physics, and why.In W.H.Zurek（editor），*Complexity, Entropy, and the Physics of*

Information. Reading, MA：Addison-Wesley，1990，PP.137—148.

Berlekamp, E.，Conway, J. H.，and Guy, R.*Winning Waysfor Your Mathematical Plays*，volume 2.San Diego, CA：Academic Press，1982.

Bickhard, M. H.The biological foundations of cognitive science.In Mind 4：*Proceedingsofthe 4th Annual Meetingofthe Cognitive Science Societyof Ireland.* Dublin, Ireland：J.Benjamins，1999.

Bolton, R. J.and Hand, D.J.Statistical fraud detection：A review.Statistical Science，17（3），2002，PP.235—255.

Bonabeau, E. Control mechanisms for distributed autonomous systems：Insights from social insects.In L.A.Segel and I.R.Cohen（editors），*Design Principlesforthe Immune System and Other Distributed Autonomous Systems.* New York：Oxford University Press，2001.

Bonabeau, E.，Dorigo, M.，and Theraulaz, G.*Swarm Intelligence*：*From Natural to Artifcial Systems.* New York：Oxford University Press，1999.

Borrell, B. Metabolic theory spat heats up.*The Scientist*（*News*），November 8，2007.（http：//www.the-scientist.com/news/display/53846/）

Bovill, C.*Fractal Geometry in Architecture and Design.* Boston：Birkh iuser. 1996.

Bowker, J.（editor）.*The Cambridge Illustrated History of Religions.* Cambridge, UK：Cambridge University Press，2002.

Bowlby, J.*Charles Darwin*：*A New Lire.* New York：Norton，1992.

Box, G. E.P.and Draper, N.R.*Empirical Model Building and Response Surfaces.* New York：Wiley，1997.

Burgard, A. P.，Nikolaev, E.V.，Schilling, C.H.，and Maranas,

C.D.Flux coupling analysis of genome-scale metabolic network reconstructions.*Genome Research* ，14，2004，PP.301—312.Burian, R.M.and Richardson, R.C.Form and order in evolutionary biology：Stuart Kauffman's transformation of the oretical biology.*In PSA*：*Proceedings of the Biennial Meeting of the Philosophy of Science Association* ，V0L.2：Symposia and Invited Papers，1 990，267—287.

Burkhardt, F. and Smith, S.（editors）.*The Correspondence of Charles Darwin* ，Volume 1.Cambridge, UK：Cambridge University Press，1985.

Burks, A. W.Von Neumann's self-reproducing automata.In A.W.Burks（editor），*Essayson Cellular Automata.* Urbana：University of Illinois Press，1970.

Calvino, I. Invisible Cities.New York：.Harcourt Brace Jovanovich，1974.（Translated by W.Weaver.）

Carlson, J. M.and Doyle, J.Complexity and robustness.*Proceedings of the National Academy of Science* ，USA，99，2002，PP.2538—2545.

Carroll, S. B.*Endless Forms Most Beautiful*：*The New Science of Evo Devo and the Making of the Animal Kingdom.* New York：Norton，2005.

Caruso, D. A challenge to gene theory, a tougher look at biotech.*New York Times* ，July 1，2007.

Churchland, P. S.，Koch, C.，and Sejnowski, T.J.What is computational neuroscience?In E.L.Schwartz（editor），*Computational Neuroscience.* Cambridge, MA：MIT Press，1 994，PP.46—55.

Clauset, A.，Shalizi, C. R.，and Newman, M.E.J.Power-law distributions in empirical data.Preprint，2007.（http：//arxiv.org/abs/0706.1062）

Coale, K. Darwin in a box.*Wired* ，June 14，1997.

Cochrane, E. Viva Lamarck：A brief history of the inheritance of

acquired characteristics.Aeon，2（2），1997，5—39.

Cohen, I. Informational landscapes in art, science, and evolution.*Bulletinof Mathematical Biology*，68，2006，pp.1213—1229.

Cohen, I. Immune system computation and the immunological homunculus.In O.Niestrasz, J.Whittle, D.Harel, and G.Reggio（editors），Mo DELS 2006，*Lecture Notesin Computer Science*，4199，2006，pp.499-512.Berlin：Springer-Verlag.

Cohen, R.，ben-Avraham, D.，and Havlin, S. Efficient immunization strategies for computer networks and populations.*Physics Review Letters*，91（24），2003，pp.247—901.

Connell, J. H.*Minimalist Mobile Robotics*：*A Colony-Style Architecture for an Artificial Creature*. San Diego, CA：Academic Press，1990.

Cook, M. Universality in elementary cellular automata.*Complex Systems*，15（1），2004，pp.1—40.

Coullet, P. and Tresser, C.Itérations d’endomorphismes et groupe de renormalization.*Comptes Rendues de Académie des Sciences*，Paris A，287，1978，pp.577-580.

Coy, P.，Woolley, S.，Spiro, L. N.，and Glasgall, W.Failed wizards of Wall Street.*Business Week*，September 21，1998.

Crutchfield, J. P.The calculi of emergence.*Physica D*，75，1994，pp.11-54.

Crutchfield, J. P.，Farmer, J.D.，Packard, N.H.，and Shaw, R.S.Chaos.*Scientifc American*，255，December 1986.

Crutchfield, J. P.，and Hanson, J.E.Turbulent pattern bases for cellular automata.*Physica D*，69，1993，pp.279—301.

Crutchfield, J. P.，Mitchell, M.，and Das, R.Evolutionary design of

collective computation in cellular automata.In J.P.Crutchfield and P.K.Schuster（editors），*Evolutionary Dynamics—Exploring the Interplay of Selection, Neutrality, Accident, and Function.* New York：Oxford University Press，2003，pp.361—411.

Crutchfield, J. P.and Shalizi, C.R.Thermodynamic depth of causal states：When paddling around in Occam's pool shallowness is a virtue.*Physical Review E*，59（1），1999，pp.275—283.

Crutchfield, J. P.and Young, K.Inferring statistical complexity.*Physical Review Letters* 63，1989，pp.105—108.

Cupitt, R. T.Target rogue behavior, not rogue states.The*Nonproliferation Review*，3，1996，pp.46—54.

Cupitt, R. T.and Grillot, S.R.COCOM is dead, long live COCOM：Persistence and change in multilateral security institutions.*British Journalof Political Science*，27，1997，pp.361-389.

Darwin, C.*The Autobiography of Charles Darwin.* Lanham, MD：Barnes&Noble Publishing，2005. Originally published 1887.

Darwin, C. and Barlow, N.D.*The Autobiographyof Charles Darwin*（Reissue edition）.New York：Norton，1993.（Originally published 1958.）

Darwin, E.*The Temple of Nature*；*or, The Origin of Society*：*A Poem with Philosophical Notes.* London：J. Johnson，1803.

Dawkins, R.*The Extended Phenotype*（Reprint edition）. New York：Oxford University Press，1989.（Originally published 1982.）

Dean, W.，Santos, F.，Stojkovic, M.，Zakhartchenko, V.，Walter, J.，Wolf, E.，and Reik, W. Conservation of methylation reprogramming in mammalian development：

Aberrant reprogramming in cloned embryos.*Proceedings of the National*

Academy of Science，USA，98（24），2001，pp. 13734—13738.

Dennett, D. R.*Elbow Room*：*The Varieties of Free Will Worth Wanting.* Cambridge, MA：MIT Press，1984.

Dennett, D. R.*Consciousness Explained.* Boston：Little, Brown&Co，1991.

Dennett, D. R.*Darwin's Dangerous Idea.* New York：Simon&Schuster，1995.

Descartes, R.*ADiscourseonthe Method.* Translated by Ian Maclean. London：Oxford World's Classics, Oxford University Press，2006（1637）.

Dietrich, M. R.The origins of the neutral theory of molecular evolution.*Journalofthe History of Biology*，27（1），1994，pp.21—59.

Dorigo, M. and Stützle, T.*Ant Colony Optimization*，MIT Press，2004.

Dover, G. A.On the edge.*Nature*，365，1993，pp.704—706.

Dukas, H. and Hoffmann B.（editors）.*Albert Einstein, The Human Side*：*New Glimpsesfrom His Archives.* Princeton University Press，1979.

Dunne, J. A.The network structure of food webs.In M.Pascual and J.A.Dunne（editors）.*Ecological Networks*：*Linking Structure to Dynamics in Food Webs.* New York：Oxford University Press，2006，pp.27—86.

Dunne, J. A.Williams, R.J.and Martinez, N.D.Food-web structure and network theory：The role of connectance and size.*Proceedingsofthe National Academyof Science*，USA，99（20），2002，pp.12917—12922.

Eddington, A. E.*The Nature of the Physical World.* New York：Macmillan，1928.

Eigen, M. How does information originate?Principles of biological self-organization.In S.A.Rice（editor），*For Ilya Prigogine*. New York：Wiley 1978，pp.211—262.

Eigen, M.*Steps Towards Life*. Oxford：Oxford University Press，1992.

Einstein, A.*Out of My Later Years* （revised edition）. Castle Books，2005.（originally published 1950，New York：Philosophical Books.）

Eldredge, N. and Tattersall, I.*The Myths of Human Evolution*. New York：Columbia University Press，1982.

Erwin, D. H.The developmental origins of animal bodyplans.In S.Xiao and A.J.Kaufman（editors），*Neoproterozoic Geobiology and Paleobiology*. New York：Springer，2006，pp.159—197.

Everdell, W. R.*The First Moderns*：*Profles in the Origins of Twentieth-Century Thought*. Chicago, IL：University of Chicago Press，1998.

Feigenbaum, M. J.Universal behavior in nonlinear systems.*Los Alamos Science*，1（1），1980，pp.4—27.

Fell, D. A.and Wagner, A.The small world of metabolism.*Nature Biotechnology*，18，2000，pp.1121—1122.

Felton, M. J.Survival of the fittest in drug design.*Modern Drug Discovery*，3（9），2000，pp.49-50.

Floridi, L.，Open problems in the philosophy of information.*Metaphilosophy*，35（4），2004，pp. 554—582.

Fogel, D. B.*Evolutionary Computation*：*The Fossil Record*. New York：Wiley-IEEE Press，1998.

Forrest, S.*Emergent Computation. Cambridge* ，MA：MIT Press，1991.

Franks, N. R.Army ants：Acollective intelligence.*American Scientist* ，

77（2），1989，pp.138—145.

Freud, S.*Moses and Monotheism*. New York：Vintage Books，1939.

Friedheim, R. L.Ocean governance at the millennium：Where we have been, where we should go：Cooperation and discord in the world economy.*Ocean and Coastal Management* ，42（9），pp.747—765，1999.

Fry, I.*The Emergenceof Lifeon Earth*：*AHistoricaland Scientifc Overview*. Piscataway, NJ：Rutgers University Press，2000.

Galan, J. M.and Izquierdo, L.R.Appearances can be deceiving：Lessons learned reimplementing Axelrod's‘Evolutionary Approaches to Norms'.*Journalof Artifcial Societiesand Social Simulation* ，8（3），2005.（http：//jasss.soc.surrey.ac.uk/8/3/2.html）

Garber, D. Descartes, mechanics, and the mechanical philosophy.*Midwest Studies in Philosophy* ，26（1），2002，pp.185—204.

Gell-Mann, M.*The Quark and the Jaguar*. New York：Freeman，1994.

Gell-Mann, M. What is complexity?*Complexity* ，1（1），1995，pp.16—19.

Gide, A.*The Counterfeiters*. Translated by D. Bussy.New York：Vintage，1973，p.353.Original：*Journal des Faux-Monnayeurs* ，Librairie Gallimard, Paris，1927.

Gladwell, M.*The Tipping Point*：*How Little Things Can Makea Big Diference*. Boston：Little, Brown，2000.

Gleick, J.*Chaos*：*Making a New Science*. New York：Viking，1987.

Gleick, J.*Isaac Newton* ，New York：Pantheon Books，2003.

Goldstine, H. H.*The Computer*, *from Pascal to von Neumann*. Princeton,

NJ：Princeton University Press，1993.First edition，1972.

Goodrich, C. S.and Matache, M.T.The stabilizing effect of noise on the dynamics of a Boolean network，*Physica* A，379（1），2007，pp.334—356.

Gordon, D. M.Task allocation in ant colonies.In L.A.Segel and I.R.Cohen（editors），*Design Principles for the Immune System and Other Distributed Autonomous Systems.* New York：Oxford University Press，2001.

Gordon, D. M.Control without hierarchy.*Nature*，446（7132），2007，p.143.

Gould, S. J.Is a new and general theory of evolution emerging? *Paleobiology*，6，1980，pp.119-130.

Gould, S. J.Sociobiology and the theory of natural selection.In G.W.Barlow and J.Silverberg（editors），*Sociobiology*：*Beyond Nature/Nurture?* pp.257—269.Boulder, CO：Westview Press Inc.，1980.

Government Accounting Office.*Long-Term Capital Management*：*Regulators Needto Focus Greater Attentionon Systemic Risk.* Report to Congressional Request，1999.（http：//www. gao.gov/cgi-bin/getrpt?GGD-00-3）

Grant, B. The powers that be.*The Scientist*，21（3），2007.

Grene, M. and Depew, D.，*The Philosophyof Biology*：*An Episodic History.* Cambridge, U.K.：Cambridge University Press，2004.

Grinnell, G. J.The rise and fall of Darwin's second theory.*Journalofthe Historyof Biology*，18（1），1985，pp.51—70.

Grosshans, H. and Filipowicz, W.The expanding world of small RNAs.*Nature*，451，2008，pp.414—416.

Hales, D. and Arteconi, S.SLACER：A Self-Organizing Protocol for Coordination in Peer-to-Peer Networks.*IEEE Intelligent Systems* ，21（2），2006，pp.29—35.

Hardin, G. The tragedy of the commons.*Science* ，162，1968，pp.1243—1248.

Heims, S.*The Cybernetics Group.* Cambridge, MA：MIT Press，1991.

Heims, S. J.*John von Neumann and Norbert Wiener*：*From Mathematics to the Technologies of Life and Death.* Cambridge：MIT Press，1980.

Hobbes, T.*Leviathan* . Cambridge, U.K.：Cambridge University Press（1651/1991）.

Hodges, A.*Alan Turing*：*The Enigma.* New York：Simon&Schuster，1983.

Hoelzer, G. A.Smith, E.，and Pepper, J.W.，On the logical relationship between natural selection and*self-organization.Journal of Evolutionary Biology*，19（6），2007，pp.1785—1794.

Hofmeyr, S. A.An interpretive introduction to the immune system.In L.A.Segel and I.R.Cohen（editors），*Design Principlesforthe Immune Systemand Other Distributed Autonomous Systems.* New York：Oxford University Press，2001.

Hofmeyr, S. A.and Forrest, S.Architecture for an artificial immune system.*Evolutionary Computation*，8（4），2000，pp.443—473.

Hofstadter, D. R.*Gödel, Escher, Bach*：*an Eternal Golden Braid.* New York：Basic Books，1979.

Hofstadter, D. R.Mathematical chaos and strange attractors.Chapter 16 in*Metamagical Themas.* New York：Basic Books，1985.

Hofstadter, D. R.The Genetic Code：Arbitrary?Chapter 27*in Metamagical Themas*. Basic Books，1985.

Hofstadter D.*Fluid Concepts and Creative Analogies*. New York：Basic Books，1995.

Hofstadter, D.*I am a Strange Loop*. New York：Basic Books，2007.

Hofstadter, D. R.and Mitchell, M.The Copycat project：A model of mental fluidity and analogy-making.In K.Holyoak and J.Barnden（editors），*Advancesin Connectionistand Neural Computation Theory*，Volume 2：Analogical Connections，1994，pp.31—112.

Holland, J. H.*Adaptation in Natural and Artifcial Systems*. Cambridge, MA：MIT Press，1992.（First edition，1975.）

Holland, J. H.*Emergence*：*From Chaos to Order*. Perseus Books，1998.

Hölldobler, B.and Wilson, E.O.*The Ants*. Cambridge, MA：Belknap Press，1990.

Holtham, C. Fear and opportunity.*Information Age*，July 11，2007.（http：//www.informationage.com/article/2006/february/fear_and_opportun

Horgan, J. From complexity to perplexity.*Scientifc American*，272，June 1995，pp.74—79.

Horgan, J.*The End of Science*：*Facing the Limits of Knowledge in the Twilight of the Scientifc Age*. Reading, MA：Addison-Wesley，1996.

Huberman, B. A.and Glance, N.S.Evolutionary games and computer simulations.*Proceedings of the National Academy of Science*，USA，90，1993，pp.7716—7718.

Hüttenhofer, A.，Scattner, P.，and Polacek, N.Non-coding RNAs：Hope or Hype?*Trends in Genetics*，21（5），2005，pp.289—297.

Huxley, J.*Evolution*：*The Modern Synthesis.* New York, London：Harper&Brothers，1942.

Jeong, H.，Tombor, B.，Albert, R.，Oltvai, Z. N.，and Barbási, A.-L.The large-scale organization of metabolic networks.*Nature* ，407，2000，pp.651—654.

Joy, B. Why the future doesn't need us.*Wired* ，April 2000.

Kadanoff, Leo P. Chaos：A view of complexity in the physical sciences.*In From Orderto Chaos*：*Essays*：*Critical, Chaotic, and Otherwise.* Singapore：World Scientific，1993.

Kauffman, S. A.*The Origins of Order*. New York：Oxford University Press，1993.

Kauffman, S. A.*At Home in the Universe.* New York：Oxford University Press，1995.

Kauffman, S. A.*Investigations.* New York：Oxford University Press，2002.

Keller, E. F.Revisiting“scale-free”networks.*Bio Essays* ，27，2005，pp.1060-1068.

Kleinfeld, J. S.Could it be a big world after all?The“six degrees of separation”myth.Society，39，2002.

Kleinfeld, J. S.Six degrees：Urban myth?*Psychology Today* ，74，March/April 2002.

Kollock, P. The production of trust in online markets.In E.J.Lawler, M.Macy, S.Thyne, and H.A.Walker（editors），*Advancesin Group Processes* ，16.Greenwich, CT：JAI Press，1999.

Kozlowski, J. and Konarzweski, M.Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant?

Functional Ecology，18，2004，pp.283—289.

Kubrin, D. Newton and the cyclical cosmos：Providence and the mechanical philosophy.*Journal of the History of Ideas*，28（3），1967.

Kurzweil, R.*The Age of Spiritual Machines*：*When Computers Exceed Human Intelligence.* New York：Viking，1999.

Langton, C. G.*Artifcial Life*：*An Overview.* Cambridge, MA：MIT Press，1997.

Laplace, P. S.*Essai Philosophique Sur Les Probabilites.* Paris：Courcier，1814.

Lee, B. and Ajjarapu, V.Period-doubling route to chaos in an electrical power system.*IEE Proceedings, Part C*，140，1993，pp.490—496.

Leff, H. S.and Rex, A.F.*Maxwell's Demon*：*Entropy, Information, Computing.* Princeton, NJ：Princeton University Press.Second edition 2003，Institute of Physics Pub.，1990.

Leibniz, G. In C.Gerhardt（Ed.），*Die Philosophischen Schriftenvon Gottfried Wilhelm Leibniz.* VolumeⅦ.Berlin：Olms，1890.

Lesley, R. Xu, Y.，Kalled, S.L.，Hess, D.M.，Schwab, S.R.，Shu, H.-B.，and Cyster, J.G.Reduced competitiveness of autoantigen-engaged B cells due to increased

dependence on BAFF.*Immunity*，20（4），2004，pp. 441—453.

Levy, S. The man who cracked the code to everything.*Wired*，Issue 10.06，June 2002.

Lewin, R.*Complexity*：*Life at the Edge of Chaos.* New York：Macmillan，1992.

Li, M. and Vitanyi, P.*An Introduction to Kolmogorov Complexity and Its Applications.* 2nd Edition.New York：Springer-Verlag，1997.

Li, T. Y.and Yorke, J.A.Period three implies chaos.*American Mathematical Monthly* 82，1975，p.985.

Liebreich, M. How to Save the Planet：Be Nice, Retaliatory, Forgiving，&Clear.White Paper, New Energy Finance, Ltd.，2007.（http：//www.newenergyfinance.com/docs/Press/NEF_WP_Carbon-Game-Theory_05.pdf）

Liljeros, F.，Edling, C. R.，Nunes Amaral, L.A.，Stanely, H.E.，and Aberg, Y.The web of human sexual contacts.*Nature*，441，2001，pp.907—908.

Lioni, A.，Sauwens, C.，Theraulaz, G.，and Deneubourg, J.-L. Chain formation in*OEcophylla longinoda.Journal of Insect Behavior*，14（5），2001，pp.679—696.

Liu, H. A brief history of the concept of chaos，1999.（http：//members.tripod.com/～huajie/Paper/chaos.html）

Lloyd, S. The calculus of intricacy.*The Sciences*，30，1990，pp.38—44.

Lloyd, S. Measures of complexity：A non-exhaustive list.*IEEEControl Systems Magazine*，August 2001.

Lloyd, S. and Pagels, H.Complexity as thermodynamic depth.*Annals of Physics*，188，1988，pp.186—213.

Locke, J.*An Essayconcerning Human Understanding.* Edited by P. H.Nidditch.Oxford：Clarendon Press，1690/1975.

Lohr, S. This boring headline is written for Google.*New York Times*，April 9，2006.

Lorenz, E. N.Deterministic nonperiodic flow.*Journal of Atmospheric Science*，357，1963，pp.130—141.

Lovelock, J. E.*The Ages of Gaia*. New York：Norton，1988.

Luisi, P. L.*The Emergenceof Life*：*from Chemical Originsto Synthetic Biology*. Cambridge, U.K.：Cambridge University Press，2006.

Mackenzie, D. Biophysics：New clues to why size equals destiny.*Science*，284（5420），1999，pp.1607-1609.

Macrae, N.*John von Neumann*. New York：Pantheon Books，1992.

Maddox, J. Slamming the door.*Nature*，417，2007，p.903.

Malone, M. S.God, Stephen Wolfram, and everything else.*Forbes ASAP*，November 27，2000.（http：//members.forbes.com/asap/2000/1127/162.html）

Mandelbrot. B.An informational theory of the statistical structure of languages.In W.Jackson（editor），*Communication Theory*，Woburn, MA：Butterworth，1953，pp.486—502.

Mandelbrot, B. B.*The Fractal Geometry of Nature*. New York：W.H.Freeman，1977.

Markoff, J. Brainy robots start stepping into daily life，*New York Times*，July 18，2006.

Marr, D.*Vision*. San Francisco：Freeman，1982.

Mattick, J. S.RNA regulation：A new genetics?*Nature Reviews*：*Genetics*，5，2004，pp.316—323.

Maturana, H. R.and Varela, F.J.*Autopoiesisand Cognition*：*The Realizationofthe Living*，Boston：D.Reidel Publishing Co.，1980.

Maxwell, J. C.*Theory of Heat*. London：Longmans, Green and Co，1871.

May, R. M.Simple mathematical models with very complicated

dynamics.*Nature*，261，459—467，1976.

Mayr, E. An overview of current evolutionary biology.*In New Perspectiveson Evolution*，1991，pp.1-14.

Mc Allister, J. W.Effective complexity as a measure of information content.*Philosophyof Science*，70，2003，pp.302—307.

Mc Clamrock, R. Marr's three leves：A re-evaluation.*Mindsand Machines*，1（2），1991，pp.185-196.

Mc Culloch, W. and Pitts, W.A logical calculus of ideas immanent in nervous activity.*Bulletin of Mathematical Biophysics*，5，1942，pp.115—133.

Mc Shea, D. W.The hierarchical structure of organisms：A scale and documentation of a trend in the maximum.*Paleobiology*，27（2），2001，pp.405—423.

Metropolis, N.，Stein, M. L.，and Stein, P.R.On finite limit sets for transformations on the unit interval.*Journal of Combinatorial Theory*，15（A），1973，pp.25—44.

Milgram, S. The small-world problem.*Psychology Today*，1，1967，pp.61—67.

Millay, E. St.Vincent.*Minethe Harvest*：*ACollectionof New Poems.* New York：Harper，1949.

Miller, G. A.Some effects of intermittent silence.*The American Journalof Psychology*，70，1957，pp.311—314.

Millonas, M. M.The importance of being noisy.*Bulletinofthe Santa Fe Institute*，Summer，1994.

Minsky, M.*The Society of Mind*，Simon&Schuster，1987.

Mitchell, M. Computation in cellular automata：A selected review.In

T.Gramss et al.（editors），*Nonstandard Computation.* Weinheim, Germany：Wiley-VCH，1998，pp.95—140.

Mitchell, M.*Analogy-Making* as*Perception*，MIT Press，1993.

Mitchell, M. Analogy-making as a complex adaptive system.In L.Segel and I.Cohen（editors），*Design Principles for the Immune System and Other Distributed Autonomous Systems.* New York：Oxford University Press，2001.

Mitchell, M. Life and evolution in computers.*Historyand Philosophyofthe Life Sciences*，23，2001，pp.361—383.

Mitchell, M. Complex systems：Network thinking.*Artifcial Intelligence*，170（18），2006，pp.1194—1212.

Mitchell, M. Crutchfield, J.P.，and Das, R.Evolving cellular automata to perform computations：A review of recent work.*In Proceedingsofthe First International Conference on Evolutionary Computation and its Applications*（*Ev CA'96*）.Moscow：Russian Academy of Sciences，1996，pp.42—55.

Mitzenmacher, M. A brief history of generative models for power law and lognormal distributions.*Internet Mathematics*，1（2），2003，pp.226—251.

Montoya, J. M.and Solé，R.V.Small world patterns in food webs.*Journalof Theoretical Biology*，214（3），2002，pp.405—412.

Moore, C. Recursion theory on the reals and continuous-time computation.*Theoretical Computer Science*，162，1996，pp.23—44.

Moravec, H.*Robot*：*Mere Machine to Transcendent Mind.* New York：Oxford University Press，1999.

Morton, O. Attack of the stuntbots.*Wired*，12.01.2004.

Mukherji, A.，Rajan, V.，and Slagle, J. R.Robustness of cooperation.*Nature*，379，1996，pp.125—126.

Muotri, A. R.，Chu, V.T.，Marchetto, M.C.N.，Deng, W.，Moran, J.V.and Gage, F.H.Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.*Nature*，435，2005，pp.903—910.

Nagel, E. and Newman, J.R.*Gödel’s Proof* .New York：New York University Press，1958.

Newman, M. E.J.Power laws, Pareto distributions and Zipf's law.Contemporary Physics，46，2005，pp.323—351.

Newman, M. E.J.，Moore, C.，and Watts, D.J.Mean-field solution of the small-world network model.Physical Review Letters，84，1999，pp.3201—3204.

Newman, M. E.J.，Forrest, S.，and Balthrop, J.Email networks and the spread of computer viruses.*Physical Review E*，66，2002，pp.035—101.

Nicolis, G. and Progogine, I.*Exploring Complexity*. New York：W.H.Freeman，1989.

Niklas, K. J.Size matters！*Trends in Ecology and Evolution*，16（8），2001，p.468.

Nowak, M. A.Five rules for the evolution of cooperation.*Science*，314（5805），2006，pp.1560-1563.

Nowak, M. A.，Bonhoeffer, S.，and May, R.M.Spatial games and the maintenance of cooperation.*Proceedings of the National Academy of Sciences*，USA，91，1994，pp.4877-4881.

Nowak, M. A.，Bonhoeffer, S.，and May, R.M.Reply to Mukherji et al.*Nature*，379，1996，p.126.

Nowak, M. A.and May, R.M.Evolutionary games and spatial

chaos.*Nature*，359（6398），1992，pp.826-829.

Nowak, M. A.and Sigmund, K.Biodiversity：Bacterial game dynamics.*Nature*，418，2002，pp.138-139.

Packard, N. H.Adaptation toward the edge of chaos.In J.A.S.Kelso, A.J.Mandell, M.F.Shlesinger（editors），*Dynamic Patternsin Complex Systems*，pp.293-301.Singapore：World Scientific，1988.

Pagels, H. The*Dreams of Reason.* New York：Simon&Schuster，1988.

Paton, R.，Bolouri, H.，Holcombe, M.，Parish, J. H.，and Tateson, R.（editors）.*Computation in Cells and Tissues*：*Perspectives and Tools of Thought.* Berlin：Springer-Verlag，2004.

Peak, D.，West, J. D.，Messinger, S.M.，and Mott, K.A.Evidence for complex, collective dynamics and emergent, distributed computation in plants.*Proceedingsofthe National Academy of Sciences*，USA，101（4），2004，pp.918-922.

Pearson, H. What is a gene?*Nature*，441，2006，pp.399-401.

Pierce, J. R.*An Introduction to Information Theory*：*Symbols, Signals, and Noise.* New York：Dover，1980.（First edition，1961.）

Pines, D.（editor）.*Emerging Synthesesin Science.* Reading, MA：Addison-Wesley，1988.

Poincaré，H.*Scienceand Method.* Translated by Francis Maitland.London：Nelson and Sons，1914.

Poundstone, W.*The Recursive Universe.* William Morrow，1984.

Poundstone, W.*Prisoner's Dilemma.* New York：Doubleday，1992.

Price, D. J.Networks of scientific papers.Science，149，1965，pp.510—515.

Provine, W. B.*The Originsof Theoretical Population Genetics*. University of Chicago Press，1971.

Rapoport, A.*General System Theory*：*Essential Concepts and Applications*，Cambridge, MA：Abacus Press，1986.

Redner, S. How popular is your paper?An empirical study of the citation distribution.*European Physical Journal B*，4（2），1998，pp.131—134.

Regis, E.*Who Got Einstein's Ofce?Eccentricity and genius at the Institute for Advanced Study*. Menlo Park, CA：Addison-Wesley，1987.

Rendell, P. Turing universality of the game of Life.In A.Adamatzky（editor），*Collision-Based Computing* .London：Springer-Verlag，2001，pp.513—539.

Robbins, K. E.，Lemey, P.，Pybus, O.G.，Jaffe, H.W.，Youngpairoj, A.S.，Brown, T.M.，Salemi, M.，Vandamme, A.M.and Kalish, M.L.，U.S.human immunodeficiency virus type 1 epidemic：Date of origin, population history, and characterization of early strains.*Journal of Virology*，77（11），2003，pp.6359-6366.

Rota, G-C, In memoriam of Stan Ulam—The barrier of meaning.*Physica D*，2，1986，pp. 1—3.

Rothman, T. The evolution of entropy.Chapter 4 in*Scienceàla Mode*. Princeton, NJ：Princeton University Press，1989.

Russell, B. A*History of Western Philosophy*，Touchstone，1967（First edition，1901）.

Schlossberg, D. LAX Computer Crash Strands International Passengers.*Consumer Afairs.com*，August 13，2007.（http：//www.consumeraffairs.com/news04/2007/08/lax_computers.html）

Schneeberger, A. Mercer, C.H.，Gregson, S.A.，Ferguson, N.M.，Nyamukapa, C.A.，Anderson, R.M.，Johnson, A.M.，and Garnett,

G.P.Scale-free networks and sexually transmitted diseases：A description of observed patterns of sexual contacts in Britain and Zimbabwe.*Sexually Transmitted Diseases*，31（6），2004，pp.380—387.

Schwartz, J. Who needs hackers?*New York Times*，September 12，2007.

Selvam, A. M.The dynamics of deterministic chaos in numerical weather prediction models.*Proceedings of the American Meteorological Society*，*8th Conference on Numerical Weather Prediction*，Baltimore, MD，1988.

Shalizi, C. Networks and Netwars，2005.Essay at（http：//www.cscs.umich.edu/～crshalizi/weblog/347.html）.

Shalizi, C. Power Law Distributions，1/f noise, Long-Memory Time Series，2007.Essay at（http：//cscs.umich.edu/～crshalizi/notebooks/power-laws.html）.

Shannon, C. A mathematical theory of communication.*The Bell System Technical Journal*，27，1948，pp.379—423，623—656.

Shaw, G. B.*Annajanska, the Bolshevik Empress.* Whitefish, MT：Kessinger Publishing，2004.（Originally published 1919.）

Shouse, B. Getting the behavior of social insects to compute.*Science*，295（5564），2002，p.2357.

Sigmund, K. On prisoners and cells.*Nature*，359（6398），1992，p.774.

Simon, H. A.On a class of skew distribution functions.*Biometrika*，42（3—4），1955，p.425.

Simon, H. A.The architecture of complexity.*Proceedingsofthe American Philosophical Society*，106（6），1962，pp.467—482.

Sompayrac, L. M.*How the Immune System Works* ，2nd edition, Blackwell Publishing，1991.

Stam, C. J.and Reijneveld, J.C.Graph theoretical analysis of complex networks in the brain.*Nonlinear Biomedical Physics* ，1（1），2007，p.3.

Stoppard, T.*Arcadia* . New York：Faber&Faber，1993.

Strogatz, S.*Nonlinear Dynamics and Chaos*. Reading, MA：Addison-Wesley，1994.

Strogatz, S.*Sync*：*How Order Emerges from Chaos in the Universe, Nature, and Daily Life.* New York：Hyperion，2004，p. 287.

Szilard, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings.*Zeitschrift fuer Physik* ，53，1929，pp.840—856.

Tattersall, I.*Becoming Human*：*Evolution and Human Uniqueness.* New York：Harvest Books，1999.

Travis, J. Eye-opening gene.*Science News Online* ，May 10，1997.

Turing, A. M.On computable numbers, with an application to the*Entscheidungsproblem.Proceedings of the London Mathematical Society* ，2（42），1936，pp.230—265.

Ulam, S. M.and von Neumann, J.On combination of stochastic and deterministic processes（abstract）.*Bulletinofthe American Mathematical Society* ，53，1947，1120.Varn, D.P.，Canright, G.S.，and Crutchfield, J.P.Discovering planar disorder in close-packed structures from X-ray diffraction：Beyond the fault model.*Physical Review* B，66，2002，

pp. 174110-1-174110-4.

Verhulst, P.-F. Recherches mathematiques sur la loi d'accroissement de la population.*Nouv.mem.de l'Academie Royale des Sci.et Belles-Lettres de*

Bruxelles，18，1845，pp.1-41.

Von Bertalanffy, L. An outline of general system theory.*The British Journal for the Philosophy of Science*，1（92），1950，pp.134—165.

Von Bertalanffy, L.*General System Theory*：*Foundations, Development, Applications*，New York：G. Braziller，1969.

Von Neumann, J.*Theory of Self-Reproducing Automata*（edited and completed by A. W.Burks）.Urbana：University of Illinois Press，1966.

Wagner, N. R.The logistic equation in random number generation.*Proceedingsofthe Thirtieth Annual Allerton Conference on Communications, Control, and Computing*，University of Illinois at Urbana-Champaign，1993，pp.922—931.

Wang, H. Reflections on*Kurt Gödel*. Cambridge, MA：MIT Press，1987.

Watts, D. J.*Six Degrees*：*The Science of a Connected Age*. New York：Norton，2003.

Watts, D. J.and Strogatz, S.H.Collective dynamics of“small world”networks.*Nature*，393，1998，pp.440—442.

Weiner, J.*The Beakofthe Finch*：*AStoryof Evolutionin Our Time*. New York：Knopf，1994.

Wiener, N.*The Human Use of Human Beings*. Boston：Houghton Mifflin，1950.

Wiener, N.*Cybernetics*：*Or the Control and Communication in the Animal and the Machine*. Cambridge, MA：MIT Press，1948.

Wiener, P. Antibiotic production in a spatially structured environment.*Ecology Letters*，3（2），2000，pp.122-130.

West, G. B.，Brown, J.H.，and Enquist, B.J.The fourth dimension of

life：Fractal geometry and allometric scaling of organisms.*Science*，284，1999，pp.1677-1679.

West, G. B.and Brown, J.H.Life's universal scaling laws.*Physics Today*，57（9），2004，p.36.

West, G. B.and Brown, J.H.The origin of allometric scaling laws in biology from genomes to ecosystems：Towards a quantitative unifying theory of biological structure and organization.*Journal of Experimental Biology*，208，2005，pp.1575-1592.

West, G. B.，Brown, J.H.，and Enquist, B.J.Yes, West, Brown and Enquist's model of allometric scaling is both mathematically correct and biologically relevant.（Reply to Kozlowski and Konarzweski，2004.）*Functional Ecology*，19，2005，pp.735—738.

Westfall, R. S.*Never at Rest*：*A Biography of Isaac Newton.* Cambridge, U.K.：Cambridge University Press，1983.

Whitfield, J. All creatures great and small.*Nature*，413，2001，pp.342—344.

Williams, F. Artificial intelligence has a small but loyal following.*Pensionsand Investments*，May 14，2001.

Williams, S. Unnatural selection.*Technology Review*，March 2，2005.

Willinger, W.，Alderson, D.，Doyle, J. C.，and Li, L.More“normal”than normal：Scaling distributions and complex systems.In R.G.Ingalls et al.，*Proceedingsofthe2004 Winter Simulation Conference*，pp.130—141.Piscataway, NJ：IEEE Press，2004.

Wolfram, S. Universality and complexity in cellular automata.*Physica D*，10，1984，pp.1—35.

Wolfram, S. Twenty problems in the theory of cellular automata.*Physica Scripta*，T9，1985，pp.170—183.

Wolfram, S.*A New Kind of Science*. Champaign, IL：Wolfram Media，2002.

Wright, R. Did the universe just happen?*Atlantic Monthly*，April 1988，pp.29—44.

Yoon, C. K.From a few genes, life's myriad shapes.*New York Times*，June 26，2007.

Yule, G. U.A mathematical theory of evolution, based on the conclusions of Dr.J.C.Willis.*Philosophical Transactions of the Royal Society of London*，Ser.B，213，1924，pp.21—87.

Ziff, E. and Rosenfield, I.Evolving evolution.*The New York Reviewof Books*，53（8），May 11，2006.

Zipf, G. K.*Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, MA：Harvard University Press，1932.

Zuse, K. Rechnender Raum.Braunschweig：Friedrich Vieweg&Sohn，1969.English translation：*Calculating Space*. Cambridge, MA：MIT Technical Translation AZT-70-164-GEMIT, Massachusetts Institute of Technology（Project MAC），02139，February 1970.

Zykov, V. Mytilinaios, E.，Adams, B.，and Lipson, H.Self-reproducing machines.*Nature*，435，2005，pp.163—164.

[1] 还原论：Hofstadter, D.R.，*Gödel, Escher, Bach*：*an Eternal Golden Braid*. New York：Basic Books，1979，p.312。

[2] “将面临的所有问题尽可能地细分”：Descartes, R.，*A Discourse on the Method* .Ian Maclean英文翻译。Oxford：Oxford University Press，1637/2006，p.17。

[3] “大部分大的基本原理似乎”：引自Horgan, J.，*The Endof Science*：*Facingthe Limitsof Knowledge in the Twilight of the Scientifc Age.* Reading, MA：Addison—Wesley，1996，p.19。

[4] “科学中涌现的综合”：会议文集已结集出版：Pines, D.，*Emerging Syntheses in Science* .Reading, MA：Addison-Wesley，1988。

[5] “致力于研究各种高度复杂和相互作用的系统”；“推动知识的统一”：G.Cowan, Plans for the future.收录于Pines, D.，*Emerging Syntheses in Science* .Reading, MA：Addison-Wesley，1988，pp.235，237。

[6] “主题为‘涌现计算’的研讨会”：会议文集已结集出版：Forrest, S.，*Emergent Computation* .Cambridge, MA：MIT Press，1991。

[7] “科学已经探索了微观和宏观世界”：Pagels, H.，The Dreams of Reason。New York：Simon&Schuster，1988，p.12。

[8] “一些思想是由简单的思想组合而成”：Locke, J.，*An Essay Concerning Human Understanding* 。P.H.Nidditch编辑。Oxford：Clarendon Press，1690/1975，p.2.12.1。

[9] “几十万只行军蚁”：这段对行军蚁习性的介绍是根据以下来源整理：Franks, N.R.，Army ants：A collective intelligence.*American Scientist* ，77（2），1989，pp.138—145；以及H lldobler, B.Wilson, E.O.，The Ants.Cambridge, MA：Belknap Press，1990。

[10] “单只行军蚁”：Franks, N.R.，Army ants：A collective intelligence.*American Scientist* ，77（2），1989，pp.138—145。

[11] “具有所谓‘集体智能’的‘超生物’”：例如，H lldobler, B.Wilson, E.O.，*The Ants* .Cambridge, MA：Belknap Press，1990，p.107。

[12] “我研究了布氏游蚁”：Franks, N.R.，Army ants：A collective intelligence.*American Scientist* ，77（2），1989，p.140。

[13] “侯世达在《哥德尔、艾舍尔、巴赫——集异璧之大成》一书中”：Hofstadter, D.R.，Ant fugue。见*Gödel, Escher, Bach*：*an Eternal Golden Braid.* New York：Basic Books，1979。

[14] “再一次一无所知”：Stoppard, T.，Arcadia.New York：Faber&Faber，1993，pp.47—48。

[15] “自然简单而自足”：引自Westfall, R.S.，*Neverat Rest*：*ABiographyof Isaac Newton* .Cambridge：Cambridge University Press，1983，p.389。

[16] “原则上就有可能预测任何时刻的情况”：Laplace, P.S.，*Essai Philosophique Sur Les Probabilites.* Paris：Courcier，1814。

[17] “物理尺度太小”：Liu, Huajie, A brief history of the concept of chaos，1999（http：//members.tripod.com/～huajie/Paper/chaos.htm）。

[18] “洛伦兹发现”：Lorenz, E.N.，Deterministic nonperiodic flow.*Journalof Atmospheric Science*，357，1963，pp.130—141。

[19] “这是一个线性系统”：有人可能会说这个并不真的是线性系统，因为群体数量随时间呈指数增长：$n_t = 2^t n_0$ 。不过这里的线性指的是n_t 到n_{t+1} 是线性映射。

[20] “逻辑斯蒂模型”：参见（http：//mathworld.wolfram.com/Logistic Equation.html）：“逻辑斯蒂方程[有时也称为费尔哈斯特模型（Verhulst model）、逻辑斯蒂模型或逻辑斯蒂增长曲线]是种群数量增长模型，最早由费尔哈斯特（Pierre Verhulst）发表（1845）。模型是时间连续的，但从连续模型得出的离散二次迭代方程也叫逻辑斯蒂方程。”逻辑斯蒂映射是表示逻辑斯蒂模型的非常有用的方式。

[21] “我不给出逻辑斯蒂模型的具体形式”：逻辑斯蒂模型如下：$n_{t+1} = (\text{出生率}-\text{死亡率})\left[kn_t - n_t^2\right]/k,$ 其中n_t 是当前这一代的种群数量，k是承载能力。让$x_t = n_t / k$, R=（出生率-死亡率），就能从中得到逻辑斯蒂映射。其中x_t 表示“承载率”：当前种群数量与最大可能的种群数量的比率。从而 $x_{t+1} = Rx_t(1-x_t)$ 因为种群数量n_t 总是介于0和k之间，所以x_t 总是介于0和1之间。

[22] “逻辑斯蒂映射”：以下文献中有对逻辑斯蒂映射的细节讨论，有普通科学教育背景的读者可以看一看：Feigenbaum, M.J.，Universal behavior in nonlinear systems.*Los Alamos Science*，1（1），1980，pp.4—27；Hofstadter, D.R.，Mathematical chaos and strange attractors.*Metamagical Themas* .New York：Basic Books，1985；Kadanoff, Leo P.，Chaos, A view of complexity in the physical sciences.*From Order to Chaos*：*Essays*：*Critical, Chaotic, and Otherwise.Singapore*：World Scientific，1993.

[23] 科普书作者都会被告知一条定律：书中每增加一个数学公式，读者就会减少一半。我也不例外——我的编辑明确告诉了我这一点。不过我还是要列出逻辑斯蒂映射的等式，因此如果你是碰到公式就要扔书的那一半读者，请跳过下一行。

[24] “1971年，数学生物学家梅在著名的《自然》杂志上发表了一篇文章”：May, R.M.，Simple mathematical models with very complicated dynamics.*Nature*，261，pp.459—467，1976。

[25] “乌拉姆、冯·诺依曼、梅特罗波利斯、保罗·斯坦和米隆·斯坦”：Ulam, S.M.，and von

Neumann, J.，*Bulletin of the American Mathematical Society* ，53，1947，p.1120；Metropolis, N.，Stein, M.L.，&Stein, P.R.，On finite limit sets for transformations on the unit interval.*Journal of Combinatorial Theory*，15（A），1973，pp.25—44。

[26] “x的值……会变成混沌”：怎么知道系统不会在迭代许多步后最终进入周期振荡呢？这可以在数学上进行证明；例如，参见Strogtaz, S.，*Nonlinear Dynamics and Chaos* .Reading, MA：Addison-Wesley，1994，pp.368—369。

[27] “被用来在计算机中生成伪随机数”：伪随机数发生器是一个确定的函数或算法，输出值的分布和相关性满足特定的统计随机性检验。这种算法在现代计算中有许多应用。乌拉姆和冯·诺依曼首先提出可以用逻辑斯蒂映射作为伪随机数发生器的基础[Ulam, S.M.，von Neumann, J.，On combination of stochastic and deterministic processes（abstract）.*Bulletinofthe American Mathematical Society* ，53，1947，p.1120]。许多学者对此进行了深入研究，例如，Wagner, N.R.，The logistic equation in random number generation.收录在*Proceedingsofthe Thirtieth Annual Allerton Conferenceon Communications, Control, and Computing*，University of Illinois at Urbana-Champaign，1993，pp.922—931。

[28] “简单的确定性方程”：May, R.M.，Simple mathematical models with very complicated dynamics.*Nature*，261，1976，pp.459—467。

[29] “最早用术语混沌……李天岩（T.Y.Li）和约克”：Li, T.Y.&Yorke, J.A.，Period three implies chaos.*American Mathematical Monthly* 82，1975，p.985。

[30] “通往混沌的倍周期之路有着悠久的历史”：对于混沌研究的有趣历史，参见Aubin, D.&Dalmedico, A.D.，Writing the History of Dynamical Systems and Chaos：*Longue Durée* and revolution, disciplines, and cultures.*Historia Mathematica* 29，2002，pp.273—339。

[31] “费根鲍姆将其引入了动力系统理论”：对费根鲍姆理论的简单解释，参见Hofstadter, D.R.，Mathematical chaos and strange attractors.*Metamagical Themas* .New York：Basic Books，1985。

[32] “一个科学家所能遇到的最好的事情”：引自Gleick, J.，*Chaos*：*Making a New Science* .New York：Viking，1987，p.189。

[33] “一些气象计算机模型”：参见Selvam, A.M..The dynamics of deterministic chaos in numerical weather prediction models.*Proceedings of the American Meteorological Society*，*8th Conference on Numerical Weather Prediction, Baltimore* ，MD，1988；以及Lee, B.&Ajjarapu, V.，Period-doubling route to chaos in an electrical power system.*IEE Proceedings*，Part C，140，1993，pp.490—496。

[34] “科雷特和特雷瑟，他们也用重正化技术研究了倍周期分叉”：Coullet, P.，&Tresser, C.，Itérations d’endomorphismes et groupe de renormalization.*Comptes Renduesde Académiedes Sciences*，Paris A，287，1978，pp.577—580。

[35] “熵增定律”：Eddington, A.E.，*The Natureofthe Physical World* .Macmillan, New York，1928，p.74。

[36] “复杂系统比简单系统更能接收、存储和利用信息”：Cohen, I.，Informational landscapes in art, science, and evolution.*Bulletin of Mathematical Biology* ，68，2006，p.1218。

[37] “进化不仅只会用DNA要把戏”：Beinhocker, E.D.，*The Origin of Wealth*：*Evolution, Complexity, and the Radical Remaking of Economics.Cambridge* ，MA：Harvard Business School Press，2006，p.12。

[38] “虽然它们的物理属性很不相同”：Gell-Mann, M.，*The Quark and the Jaguar* .New York：W.H.Freeman，1994，p.21。

[39] “为什么第二定律能区分过去和现在”：Rothman, T.，The evolution of entropy.*Science à la Mode* .Princeton, NJ：Princeton University Press，1989，p.82。

[40] “热系统（左边）变得更热”：Maxwell，引自Leff, H.S.&Rex, A.F.，*Maxwell’s Demon*：*Entropy, Information, Computing.* Princeton University Press.Second edition 2003，Institute of Physics Pub.，1990，p.5。

[41] 一篇……著名论文”：Szilard, L.，On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings.*Zeitschrift fuer Physik* ，53，1929，pp.840—856。

[42] “数学家班尼特证明”：班尼特的论证很巧妙；细节见Bennett, C.H.，The thermodynamics of computation—a review.*International Journal of Theoretical Physics* ，21，1982，pp.905—940。这些思想中许多又被物理学家奥利弗·彭罗斯（Oliver Penrose）独立发现（Leff, H.S.and Rex, A.F.，*Maxwell’s Demon*：*Entropy, Information, Computing*， Taylor&Francis，1990；second edition, Institute of Physics Pub.，2003）。

[43] “小妖的问题依然存在争议”：例如，参见Maddox, J.，Slamming the door.Nature，417，2007，p.903。

[44] “大部分物理学家都只接受”：有证据表明玻尔兹曼自己对别人的工作也不留情面。据艾维戴尔（William Everdell）说，玻尔兹曼写了文章，题为“论叔本华（Schopenhauer）”的一篇论文，而后来又写道，他想用这样的标题，“证明叔本华是一个堕落、没有思想、无知、胡说八道的哲学家，他的观点完全是空洞的言语垃圾”。Everdell, W.R.，*The First Moderns*：*Proflesinthe Originsof Twentieth-Century Thought.* Chicago, IL：University of Chicago Press，1998，p.370。

[45] “玻尔兹曼将宏观状态的熵定义为”：玻尔兹曼熵的定义假定对应于给定的宏观态的所有微观态都有相等的概率。玻尔兹曼还给出了一个更广义的公式定义非等概率微观态的熵。

[46] “玻尔兹曼熵的公式”：玻尔兹曼熵的公式为：S=klog W，S是熵，W是对应给定宏观态的可

能微观态的数量，k是“玻尔兹曼常数”，这个数用来将熵变成标准单位。

[47] “1948年，香农发表了论文‘通信的数学理论’”：Shannon，C.，A mathematical theory of communication.*The Bell System Technical Journal*，27，1948，pp.379—423，623—656。

[48] “让通信理论和物理学联姻的努力”：Pierce，J.R.，*An Introduction to Information Theory*：*Symbols, Signals, and Noise*.New York：Dover，1980，p.24。（1961年第一版）

[49] “Quo facto”：Leibniz，G.（1890）.收录在C.Gerhardt（编辑），*Die Philosophischen Schriftenvon Gottfried Wilheml Liebniz*，VolumeⅦ.Berlin：Olms.英文引文引自Russell，B.，A History of Western Philosophy, Touchstone，1967，p.592。（1901年第一版）

[50] “细胞和组织中的计算”：E.g.，Paton，R.，Bolouri，H.，Holcombe，M.，Parish，J.H.，&Tateson.R.编辑，*Computationin Cellsand Tissues*：*Perspectivesand Toolsof Thought*，Berlin：Springer-Verlag，2004。

[51] “免疫系统的计算”：Cohen，I.R.，Immune system computation and the immunological homunculus.收录在O.Nierstrasz等（编辑），*Mo DELS2006*，*Lecture Notesin Computer Science* 4199.Springer-Verlag，2006，pp.499—512。

[52] “市场的分布式计算的本质和局限”：David Pennock题为“Information and complexity in securities markets”的演讲，Institute for Computational and Mathematical Engineering, Stanford University, November 2005。

[53] “植物中的涌现计算”：Peak，D.，West，J.D.，Messinger，S.M.，&Mott，K.A.，Evidence for complex, collective dynamics and emergent, distributed computation in plants.*Proceedingsof the National Academy of Sciences*，USA，101（4），2004，pp.918—922。

[54] “不存在不可解的问题”：引自Hodges，A.，*Alan Turing*：*The Enigma*，New York：Simon&Schuster，1983，p.92。

[55] “哥德尔的证明很复杂”：对这个证明的出色阐释，参见Nagel，E.&Newman，J.R.，*Gödel’s Proof*.New York：New York University，1958；以及Hofstadter, D.R.，*Gödel, Escher, Bach*：*an Eternal Golden Braid.* New York：Basic Books，1979。

[56] “这是在研究中惊人的转折”：Hodges，A.，*Alan Turing*：*The Enigma*.New York：Simon&Schuster，1983，p.92。

[57] “第三问题又被英国数学家图灵干掉了”：还有一位数学家丘奇（Alonzo Church），也证明了在数学中存在不可判定命题，但后来图灵的研究更有影响力。

[58] “同样，他的答案也是‘否’”：Turing，A.M.，On computable numbers, with an application to the*Entscheidungsproblem.Proceedings of the London Mathematical Society*，2（42），1936，

pp.230—265。

[59] “据他的传记作者王浩说”：Wang，H.，*Reflections on Kurt Gödel* .Cambridge，MA：MIT Press，1987。

[60] “一切伟大的真理开始时都是大逆不道”：Shaw，G.B.，*Annajanska, the Bolshevik Empress.* London：Kessinger Publishing，1919/2004，p.15。

[61] “一切行无常，生者必有尽”：引自Bowker, J.（编辑），*The Cambridge Illustrated History of Religions.* Cambridge, UK：Cambridge University Press，2002，p.76。

[62] “地必如衣服渐渐旧了”：Isaiah 51：6.*The Holy Bible, King James Version* 。

[63] “哦，夏日的芳香怎能抵抗”：引自Shakespeare, Sonnet 65。

[64] “如果要我选择一个历史上最重要的思想”：Dennett，D.R.，*Darwin’s Dangerous Idea.* New York：Simon&Schuster，1995，p.21。

[65] “无垠波涛下的有机生命”：Darwin，E.，*The Temple of Nature* ；或*The Origin of Society*：*A Poem with Philosophical Notes.* London：J.Johnson，1803。

[66] “拉马克……有一些很确凿的证据”：引自Grinnell, G.J.，The rise and fall of Darwin’s second theory.*Journal of the History of Biology* ，18（1），1985，p.53。

[67] “如果动物的本能生活能有任何解释”：Freud，S.，*Moses and Monotheism* .New York：Vintage Books，1939，pp.128—129.引自Cochrane，E.，Viva Lamarck：A brief history of the inheritance of acquired characteristics.Aeon 2：2，1997，pp.5—39。

[68] “除了打猎、养狗、抓老鼠，你还会什么”：Darwin, C.&Barlow, N.D.，*The Autobiography of Charles Darwin.* Reissue edition, New York：W.W.Norton，1958/1993，p.28。

[69] “达尔文还读了亚当·斯密的自由市场圣经”：达尔文在给W.D.Fox的一封信中写到（1829年1月），“我的研究中包括亚当·斯密和洛克（Locke）。”*The Correspondence of Charles Darwin* ，第1卷（F.Burkhardt&S.Smith编辑）.Cambridge, U.K.：Cambridge University Press，1985，p.72。

[70] “喙的形态是对不同食物来源的适应”：对此的进一步阅读，参见Weiner, J.，*The Beak of the Finch*：*A Story of Evolution in Our Time.* New York：Knopf，1994。

[71] “我深信物种不是不可改变的”：引自Bowlby, J.，*Charles Darwin*：*A New Life* .New York W.W.Norton，1992，p.254。

[72] “柏拉图……在《斐多篇》中说”：Barrett, P.（编辑），*Charles Darwin’s Notebooks* ，1836—1844：*Geology, Transmutationof Species, Metaphysical Enquiries.* Ithaca, NY：Cornell University

Press，1987，p.551。

[73] “我的所有成果”：参见达尔文1858年6月18日给莱尔的信。收录在*The Darwin Correspondence Online Database*（http：//www.darwinproject.ac.uk），Letter 2285。

[74] “有些可鄙”：参见达尔文1858年6月25日给莱尔的信。同上，Letter 2294。

[75] “我完全承认马修先生多年前”：引自Darwin，C.，*The Autobiography of Charles Darwin.* Lanham, MD：Barnes&Noble Publishing，尾注21，2005，p.382。1887年首次发表。

[76] “真蠢，我怎么没有想到！”：引自Provine，W.B.，*The Origins of Theoretical Population Genetics.* Chicago：University of Chicago Press，1971，p.4。

[77] “达尔文理论和孟德尔遗传学”：“现代综合”一词来自Julian Huxley的名著*Evolution*：*The Modern Synthesis*，New York，London：Harper，1942。现代综合也称新达尔文综合（*Neo-Darwinian Synthesis*），现代进化论综合（*Modern Evolutionary Synthesis*），熟悉的人就直接称之为综合（*Synthesis*）。

[78] “没有人在审视进化过程时能忽略现代综合的体系”：Tattersall，I.，*Becoming Human*：*Evolution and Human Uniqueness.* New York：Harvest Books，1999，p.83。

[79] “木村根据对蛋白质进化的观察提出了‘中性进化’的理论”：对木村的理论的讨论，参见Dietrich，M.R.，The origins of the neutral theory of molecular evolution.*Journalofthe Historyof Biology*，27（1），1994，pp.21—59。

[80] “艾根和舒斯特在RNA构成的病毒的进化中观察到了间断平衡现象”：对艾根和舒斯特的研究很好的阐释，参见Eigen，M.，*Steps Towards Life*.Oxford：Oxford University Press，1992。

[81] “综合理论事实上已经死了”：Gould，S.J.，Is a new and general theory of evolution emerging? *Paleobiology*，6，1980，p.120。

[82] “将进化归因于现代综合是‘20世纪生物学最大的神话’”：Eldredge，N.&Tattersall，I.，*The Myths of Human Evolution.* New York：Columbia University Press，1982，p.43。

[83] “我认为进化综合的成果并没有严重错误”：Mayr，E.，An overview of current evolutionary biology。收录在Warren，L.&Koprowski，H.（编辑），*New Perspectives on Evolution.* New York：Wiley-Liss，1991，p.12。

[84] “累积式的自然选择进化论”：Dawkins，R.，*The Extended Phenotype*（重印版）。Oxford University Press，1989，p.317。1982年首次发表。

[85] “细胞的闪存卡”：Hofstadter，D.R.The Genetic Code：Arbitrary?收录在*Metamagical Themas*.New York：Basic Books，1985，p.681。

[86] “对于如何在分子层面上定义基因的概念”：参见Pearson, H.，“What is a gene？”*Nature*，vol.441，2006，pp.399—401。

[87] “2001年，物理学家劳埃德发表了一篇文章”：Lloyd, S.，Measures of complexity：A non-exhaustive list.*IEEE Control Systems Magazine*，August 2001。

[88] 这被称为事物的算法信息量”：对柯尔莫哥洛夫、查汀和索罗蒙洛夫的思想的详细论述，参见Li, M.&Vitanyi, P.，*An Introduction to Kolmogorov Complexity and Its Applications*，2nd Edition.New York：Springer-Verlag，1997。

[89] “盖尔曼提出了一种称为‘有效复杂性’的相关度量”：Gell-Mann, M.What is complexity? *Complexity*，1（1），1995，pp.16—19。

[90] “这个定义中的主观性也有待解决”：参见，Mc Allister, J.W.，Effective complexity as a measure of information content.*Philosophy of Science* 70，2003，pp.302—307。

[91] “有逻辑深度的事物”：Bennett, C.H.，How to define complexity in physics, and why.收录在W.H.Zurek（编辑），*Complexity, Entropy, and the Physics of Information*，Reading, MA：Addison-Wesley，1990，p.142。

[92] “用最合理的方法生成某个事物时需要处理的信息量”：Lloyd, S.，The calculus of intricacy.*The Sciences*，30，1990，p.42。

[93] “劳埃德和裴杰斯提出了一种新的复杂性度量”：Lloyd, S.&Pagels, H.，Complexity as thermodynamic depth.*Annals of Physics*，188，1988，pp.186—213。

[94] “产生出这个事物最科学合理的确定事件序列”和“物理构造过程所需的热力源和信息源的总量”：Lloyd, S.，The calculus of intricacy.*The Sciences*，30，1990，p.42。

[95] “也有批评意见指出”：Crutchfield, J.P.&Shalizi, C.R.，Thermodynamic depth of causal states：When paddling around in Occam’s pool shallowness is a virtue.*Physical Review E*，59（1），1999，pp.275—283。

[96] “像物理学家沃尔夫勒姆就提出”：Wolfram, S.，Universality and complexity in cellular automata.*Physica* D，10，1984，pp.1—35。

[97] “不过，班尼特等人则认为”：例如，参见Bennett, C.H.，Dissipation, information, computational complexity and the definition of organization.收录在D.Pines（编辑），*Emerging Synthesesin Science*.Redwood City, CA：Addison-Wesley，1985，pp.215—233。

[98] “统计复杂性”：Crutchfield, J.P.&Young, K.，Inferring statistical complexity，*Physics Review Letters* 63，1989，pp.105—108。

[99] “复杂晶体的原子结构”：Varn, D.P.，Canright, G.S.，&Crutchfield, J.P.，Discovering planar disorder in close-packed structures from X-ray diffraction：Beyond the fault model.*Physical Review* B，66，2002，pp.174110-1-174110-4。

[100] “神经元的激发模式”：Haslinger, R.，Klinkner, K.L.&Shalizi, C.R.，The computational structure of spike trains.未发表手稿，2007。

[101] “宇宙也是分形的”：Mandelbrot, B.B.，*The Fractal Geometry of Nature* .New York：W.H.Freeman，1977。

[102] “一般来说分形指的是‘在任何尺度上都有微细结构’的几何形状”：Strogatz, S.，*Nonlinear Dynamics and Chaos.* Reading, MA：Addison-Wesley，1994。

[103] “这里我们直接给出结果”：对于科赫曲线，$3^{维数}$=4。为了得出维数，将两边取对数（基可为任意数）：log（$3^{维数}$）=维数×log（3）=log（4）从而维数=log（4）/log（3）≈1.26

[104] “分形维数”：对分形和分形维数概念的精彩阐述参见曼德布罗特的书——*The Fractal Geometry of Nature.* New York：W.H.Freeman，1977。

[105] “细节的瀑流”：Bovill, C.，*Fractal Geometry in Architecture and Design* .Birkhäuser Boston，1996，p.4。

[106] “《复杂性的结构》”：Simon, H.A.，The architecture of complexity.*Proceedingsofthe American Philosophical Society*，106（96），1962，pp.467—482。

[107] “复杂系统由子系统组成”：同上，p.468。

[108] “麦克西……提出了一种层次标度”：Mc Shea, D.W.，The hierarchical structure of organisms：A scale and documentation of a trend in the maximum.*Paleobiology*，27（2），2001，pp.405—423。

[109] “大自然逐渐将无生命的东西”：引自Grene, M.&Depew, D.，*The Philosophy of Biology*：*An Episodic History* .Cambridge University Press，2004，p.14。

[110] “我们都本能地知道生命是什么”：Lovelock, J.E，*The Agesof Gaia.* New York：W.W.Norton，1988，p.16。

[111] “自我复制的计算机程序”：这一章部分是来自Mitchell, M.，Life and evolution in computers.*History and Philosophy of the Life Sciences*，23，2001，pp.361—383。

[112] “……有一些让人着迷的研究”：参见，Luisi, P.L.，*The Emergenceof Life*：*From Chemical Origins to Synthetic Biology.* Cambridge：Cambridge University Press，2006；或Fry, I.，*The*

Emergence of Life on Earth：*A Historical and Scientific Overview*. Piscataway, NJ：Rutgers University Press，2000。

[113]　译注：石人是希伯莱传说中用黏土、石头或青铜制成的无生命的巨人，注入魔力后可行动，但无思考能力。皮格马利翁是奥维德《变形记》中的人物，他爱上了自己用象牙雕刻的美丽少女，并感动了爱神阿芙罗狄娜，爱神让雕塑变成了真人。

[114] “人工生命领域中”：关于人工生命领域的更多信息，参见Langton, C.G.，*Artifcial Life*：*An Overview.Cambridge*， MA：MIT Press，1997；或Adami, C.，*Introduction to Artifcial Life*. New York：Springer，1998。

[115] “1966年，巴克斯将……出版”：Von Neumann, J.，*Theory of Self-Reproducing Automata*（由巴克斯编辑补充）。Urbana：University of Illinois Press，1966。对冯·诺依曼自复制自动机的描述，参见Burks, A.W.，Von Neumann’s self—reproducing automata。收录在A.W.Burks（编辑），*Essays on Cellular Automata* .Urbana：University of Illinois Press，1970；或Mitchell, M.，Computation in cellular automata：A selected review。收录在T.Gramss等（编辑），*Nonstandard Computation*，1998，pp.95—140.Weinheim, Germany：Wiley—VCH。对DNA自我复制以及其与数理逻辑和自我复制计算机程序的关系的论述，参见Hofstadter, D.R.，Gödel, Escher, Bach：an Eternal Golden Braid.New York：Basic Books，1979，pp.495—548。

[116] “未来这种自复制机器的可能性”：引自Heims, S.J.，*John Von Neumann and Norbert Wiener*：*From Mathematics to the Technologies of Life and Death*. Cambridge, MA：MIT Press，1980，pp.212—213。

[117] “他们的书并非虚构”：Kurzweil, R.，*The Ageof Spiritual Machines*：*When Computers Exceed Human Intelligence*. New York：Viking，1999；以及Moravec, H.，Robot：*Mere Machine to Transcendent Mind*. New York：Oxford University Press，1999。

[118] “在《连线》杂志上发表了一篇后来很有名的文章”：Joy, B.，Why the future doesn’t need us.*Wired*，2000年4月。

[119] “一些简单的自复制机器人”：Zykov, V.Mytilinaios, E.，Adams, B.，&Lipson, H.，Self-reproducing machines.*Nature*，435，2005，pp.163—164。

[120] “6岁时，有一次他母亲在他面前显得心不在焉”：Macrae, N.，*John von Neumann* .New York：Pantheon，1992，p.52。

[121] “有史以来最好的数理经济学论文”：引自Macrae, N.，*John von Neumann*. New York：Pantheon，1992，p.23。

[122] “计算和计算机领域有史以来最重要的文献”：Goldstine, H.H.，*The Computer, from Pascal to von Neumann*. Princeton, NJ：Princeton University Press，第一版，1972，p.191。

[123] “匈牙利6位诺贝尔奖获得者有5位”：Mac Rae, N.，*John von Neumann* .New York：Pantheon，1992，p.32。

[124] “（IAS）数学院”：引自Macrae, N.，*John von Neumann* .New York：Pantheon，1992，p.324。

[125] “让研究院没有任何实验科学”：引自Regis, E.，*Who Got Einstein’s Office?Eccentricity and Genius at the Institute for Advanced Study*. Menlo Park, CA：Addison-Wesley，1987，p.114。

[126] “自大狂的报复”：Regis, E.，*Who Got Einstein’s Office?Eccentricity and Genius at the Institute for Advanced Study*. Menlo Park, CA：Addison-Wesley，1987，p.114。

[127] “进化计算”：对进化计算的早期研究史，参见Fogel, D.B.，*Evolutionary Computation*：*The Fossil Record*. New York：Wiley—IEEE Press，1998。

[128] “这就是遗传算法的由来”：霍兰德，引自Williams, S.Unnatural selection.*Technology Review* ，2005年3月2日。

[129] “飞行器的部分自动化设计”：Hammond, W.E.*Design Methodologies for Space Transportation Systems* ，2001，Reston, VA：American Institute of Aeronautics and Astronautics, Inc.，p.548。

[130] “分析卫星图像”：参见，Harvey, N.R.，Theiler, J.，Brumby, S.P.，Perkins, S.Szymanski, J.J.，Bloch, J.J.，Porter, R.B.，Galassi, M.，&Young, A.C.Comparison of GENIE and conventional supervised classifiers for mulitspectral image feature extraction.*IEEE Transactions on Geoscience and Remote Sensing* ，40，2002，pp.393—404.

[131] “自动化生产线的调度”：Begley, S.Software au naturel.*Newsweek* ，1995年5月8日。

[132] “设计计算机芯片”：同上。

[133] “生成逼真的动画马匹”：参见Morton, O.，Attack of the stuntbots.*Wired* ，2004年12月1日

[134] “生成逼真的演员替身动画特效”：“Virtual Stuntmen Debut in Hollywood Epic*Troy* ”新闻稿，Natural Motion Ltd（http：//www.naturalmotion.com/files/nm_troy.pdf）。

[135] “发现新药”：参见，Felton, M.J.，Survival of the fittest in drug design.*Modern Drug Discovery* ，3（9），2000，pp.49—50。

[136] “识别交易欺诈”：Bolton, R.J.&Hand, D.J.，Statistical fraud detection：A review.*Statistical Science* ，17（3），2002，pp.235—255。

[137] “分析信用卡数据”：Holtham, C.，Fear and opportunity.*Information Age* ，2007年7月11日。

[138] “预测金融市场”：参见，Williams, F.，Artificial intelligence has a small but loyal

following.*Pensions and Investments*， 2001年5月14日。

[139] “优化证券投资组合”：Coale, K.，Darwin in a box.*Wired*， 1997年6月14日。

[140] “互动式遗传算法创造的艺术作品”：参见http：//www.karlsims.com。

[141] “一个更详细的简单例子”：这个例子是受MIT人工智能实验室的一个项目的启发，那个项目中一个名为“希尔伯特”的机器人在走廊和办公室四处收集空易拉罐，并将它们送到垃圾箱。参见Connell, J.H.，*Minimalist Mobile Robotics*：*A Colony-Style Architecture for an Artificial Creature*. San Diego：Academic Press，1990。

[142] “这样就有243种可能情形”：有5个格子，每个格子中的可能情况有3种，这样就有3×3×3×3×3=243种可能情形。

[143] “进化算法是探索设计死角的伟大工具”：Jason Lohn，引自Williams, S.，Unnatural selection.*Technology Review*， 2005年3月2日。

[144] “计算机科学的真义”：引自Lewin, R.，*Complexity*：*Life at the Edge of Chaos*. New York：Macmillan，1992，p.48。

[145] “《科学》杂志不久前出版了一篇文章”：Shouse, B.，Getting the behavior of social insects to compute.*Science*， 295（5564）， 2002，2357。

[146] “‘大脑是计算机吗？’”：Churchland, P.S.，Koch, C.，&Sejnowski, T.J.，What is computational neuroscience?收录在E.L.Schwartz（编辑），*Computational Neuroscience*. Cambridge, MA：MIT Press，1994，pp.46—55。

[147] “答案是‘2的512次幂’（2^{512}）”：这一章后面会讲到，要定义一条规则，你必须说明，对于邻域内灯泡的所有可能状态组合，中心位置的灯泡下一步的状态是什么。邻域包括周围8个灯泡和中间的灯泡本身，每个灯泡的状态都可以是开或关，因此可能的状态组合的数量为$2^9=512$。而对每种状态组合，都可以用“开”或“关”作为中间灯泡下一步的状态，因此对全部512种组合可能给出的规则配置数量就为$2^{512}\approx 1.3\times 10^{154}$。

[148] “生命游戏”：这里介绍的许多内容可以参考以下文献：Berlekamp, E.，Conway, J.H.，&Guy, R.，*Winning Ways for Your Mathematical Plays*， Volume 2.San Diego：Academic Press，1982；Poundstone, W.，*The Recursive Universe*. William Morrow，1984；以及数以千计关于人工生命的网站。

[149] “康威还给出了生命游戏模拟通用计算机的证明框架”：Berlekamp, E.，Conway, J.H.，&Guy, R.，*Winning Ways for Your Mathematical Plays*， Volume 2.San Diego：Academic Press，1982。

[150] “后来由其他人细化”：例如，参见Rendell, P.，Turing universality of the game of Life。收录在A.Adamatzky（编辑），*Collision-Based Computing* ，pp.513—539.London：Springer-Verlag，2001。

[151] “注释中介绍了如何将二进制数转换为十进制数”：十进制（基为10）数的每一位对应一个10的幂，例如：$235=2\times10^2+3\times10^1+5\times10^0$ （其中$10^0=1$）。基为2时，每一位对应的则是2的幂。例如235表示成基为2的数就是11101011：$11101011=1\times2^7+1\times2^6+1\times2^5+0\times2^4+1\times2^3+0\times2^2+1\times2^1+1\times2^0=235$

[152] “规则30自动机是我在科学中所遇到的最让人惊异的事物”：引自Malone, M.S.，God, Stephen Wolfram, and everything else.*Forbes ASAP* ，2000年11月27日。（http：//members.forbes.com/asap/2000/1127/162.html）

[153] “沃尔夫勒姆对规则30印象非常深刻”：“随机序列发生器”美国专利号4691291，1987年9月1日。

[154] “类型4是有序与随机的混合”：Wolfram, S.，*A New Kind of Science* .Champaign, IL, Wolfram Media，2002，p.235。

[155] “库克最终证明了规则110的确是通用的”：Cook, M.，Universality in elementary cellular automata.*Complex Systems* 15（1），2004，1—40。

[156] “《一种新科学》……视为‘新的自然定律’”：Wolfram, S.，*A New Kind of Science.* Champaign；IL：Wolfram Media，2002，p.235。

[157] “你就能用其来解决停机问题”：参见Moore, C.，Recursion theory on the reals and continuous-time computation.*Theoretical Computer Science* ，162，1996，pp.23—44。

[158] “宇宙的终极确定性模型”：Wolfram, S.，*ANew Kindof Science.* Champaign, IL：Wolfram Media，2002，p.466。

[159] “我猜其实相当短”：沃尔夫勒姆，引自Levy, S.，The man who cracked the code to everything……*Wired* ，Issue 10.06，2002年6月。

[160] “祖斯和弗瑞德金早在20世纪60年代就提出了宇宙是元胞自动机的理论”：参见Zuse, K.，Rechnender Raum Braunschweig：Friedrich Vieweg&Sohn，1969[英文版：*Calculating Space.* MIT Technical Translation AZT-70-164-GEMIT, Massachusetts Institute of Technology（Project MAC），Cambridge, MA，02139，1970年2月]；以及Wright, R.，Did the universe just happen? *Atlantic Monthly* ，1988年4月，pp.29—44。

[161] “粒子计算”：我们在元胞自动机和粒子计算方面的研究的详细阐述可以参见Crutchfield,

J.P.，Mitchell, M.，&Das, R.，Evolutionary design of collective computation in cellular automata。收录在J.P.Crutchfield&P.K.Schuster（编辑），*Evolutionary Dynamics——Exploringthe Interplayof Selection, Neutrality, Accident, and Function.* New York：Oxford University Press，2003，pp.361—411。

[162] “物理学家帕卡德的一篇文章”：Packard, N.H.，Adaptation toward the edge of chaos。收录在J.A.S.Kelso, A.J.Mandell, M.F.Shlesinger（编辑），*Dynamic Patternsin Complex Systems.* Singapore：World Scientific，1988，pp.293—301。

[163] “多数分类”：多数分类任务在元胞自动机文献中也称为“密度分类（density classification）”。

[164] “克鲁奇菲尔德之前发明了一种方法”：参见，Crutchfield, J.P.，&Hanson, J.E.，Turbulent pattern bases for cellular automata.Physica D 69，1993，pp.279—301。

[165] “《元胞自动机理论的二十个问题》”：Wolfram, S.，Twenty problems in the theory of cellular automata.*Physica Scripta*，T9，1985，pp.170—183。

[166] “植物学家莫特、物理学家皮克和他们的同事长期观察叶子气孔的开合模式”：参见Peak, D.，West, J.D.，Messinger, S.M.，&Mott, K.A.，Evidence for complex, collective dynamics and emergent, distributed computation in plants.Proceedings of the National Academy of Sciences, USA，101（4），2004，pp.918—922。

[167] “生命系统中的信息处理”：本章部分内容来自Mitchell, M.，Complex systems：Network thinking.Artificial Intelligence，170（18），2006，pp.1194—1212。

[168] “我们就面临着以下问题”：我的问题涉及对信息处理进行描述的三个层面，这三个层面是由玛尔（David Marr）提出的：Marr, D.Vision.San Francisco, Freeman，1982。[提出类似问题的还有迈克柯蓝洛克（Ron Mc Clamrock）；参见Mc Clamrock, R.，Marr’s three levels：A re-evaluation.*Minds and Machines*，1（2），1991，pp.185—196。]

[169] “免疫系统”：对免疫系统有两个可读性很强的综述，Sompayrac, L.M.，*How the Immune System Works*，2nd edition, Blackwell Publishing，1991；以及Hofmeyr, S.A.，An interpretive introduction to the immune system.收录在L.A.Segel&I.R.Cohen（编辑），*Design Principlesforthe Immune Systemand Other Distributed Autonomous Systems.* New York：Oxford University Press，2001。

[170] “还有一种可能的机制是B细胞之间对一种有限资源的竞争”：更多细节参见Lesley, R.Xu, Y.，Kalled, S.L.，Hess, D.M.，Schwab, S.R.，Shu, H.-B.，&Cyster, J.G.，Reduced competitiveness of autoantigen-engaged B cells due to increased dependence on BAFF.*Immunity*，20（4），2004，pp.441—453。

[171] “大部分蚂蚁种类的食物搜索大致是这样进行的”：关于蚂蚁搜寻食物的更多细节，参见Bonabeau, E.，Dorigo, M.，&Theraulaz, G.，*Swarm Intelligence*：*From Natural to Artificial Systems.* New York：Oxford University Press，1999。

[172] “生态学家戈登曾研究过红色收割蚁的任务分配”：参见，Gordon, D.M.，Task allocation in ant colonies.收录在L.A.Segel&I.R.Cohen（编辑），*Design Principles for the Immune System and Other Distributed Autonomous Systems.* New York：Oxford University Press.，2001。

[173] “蚂蚁用身体搭桥和构建庇护所”：例如，参见Lioni, A.，Sauwens, C.，Theraulaz, G.，&Deneubourg, J.—L.，Chain formation in*OEcophylla longinoda.Journal of Insect Behavior* ，14（5），2001，pp.679—696。

[174] “前面描述的三个系统本质上都利用了随机性和或然性”：对随机性在复杂适应系统中的作用的探讨还可参见Millonas, M.M.，The importance of being noisy.*Bulletin of the Santa Fe Institute* ，Summer，1994。

[175] “最终蚂蚁会建立通往食物源的路径的详细地图”：Ziff, E.&Rosenfield, I.，Evolving evolution.*The New York Review of Books* ，53，8，2006年5月11日。

[176] “‘并行级差扫描’”：Hofstadter D.，*Fluid Conceptsand Creative Analogies* .New York：Basic Books，1995，p.92。

[177] “在两种探测模式中保持适当的平衡都是关键”：讨论参见Holland, J.H.，*Adaptation in Natural and Artificial Systems.* Cambridge, MA：MIT Press，1992（1975年第1版）；以及Hofstadter, D.R.&Mitchell, M.，The Copycat project：A model of mental fluidity and analogy-making.收录在K.Holyoak&J.Barnden（编辑），*Advances in Connectionist and Neural Computation Theory, Volume 2*：*Analogical Connections* ，1994，pp.31—112。

[178] “那么是谁或是什么在觉察当前情势的意义”：许多书和文章从哲学立场探讨了这个问题，其中包括：Hofstadter, D.R.，*Gödel, Escher, Bach*：*an Eternal Golden Braid.* New York：Basic Books，1979；Dennett, D.R.，Consciousness Explained.Boston：Little, Brown，1991；Bickhard, M.H.，The biological foundations of cognitive science.收录在*Mind 4*：*Proceedingsofthe4th Annual Meetingofthe Cognitive Science Society of Ireland.* Dublin, Ireland：J.Benjamins，1999；Floridi, L.，Open problems in the philosophy of information.Metaphilosophy，35（4），2004，pp.554—582；以及Hofstadter, D.，*Iam a Strange Loop.* New York：Basic Books，2007。

[179] “人工免疫系统”：参见，Hofmeyr, S.A.&Forrest, S.，Architecture for an artificial immune system.*Evolutionary Computation* ，8（4），2000，pp.443—473。

[180] “蚁群算法”：参见，Dorigo, M.&Stützle, T.，*Ant Colony Optimization* ，MIT Press，2004。

[181] 这一章部分内容来自Mitchell, M.，*Analogy-Making as Perception* ，MIT Press，1993；以及

Mitchell, M.，Analogy-making as a complex adaptive system.收录在L.Segel&I.Cohen（编辑），*Design Principles for the Immune System and Other Distributed Autonomous Systems.* New York：Oxford University Press，2001。

[182] “一年前，《萨克拉门托蜂报》”：Lohr, S.，This boring headline is written for Google.*New York Times*， April 9，2006。

[183] “在会议上你会听到人们说‘人类级的AI’”：2006年7月18日霍维茨，引自Markoff, J.，Brainy robots start stepping into daily life.*New York Times*， July 18，2006。

[184] “容易的事很难”：Minsky, M.，*The Society of Mind* .New York：Simon&Schuster，1987，p.29。

[185] 所有对真理的认识都是通过类比得来的”：Thoreau, H.D.（&L.D.Walls，编辑）.*Material Faith*：*Thoreau on Science.* New York：Mariner Books，1999，p.28。

[186] “一本新出的书，作者是印第安纳大学的一位计算机系教授”：Hofstadter, D.R.，*Gödel, Escher, Bach*：*an Eternal Golden Braid.* New York：Basic Books，1979。

[187] “由侯世达的研究团队研发的一系列让人印象更加深刻的程序”：对这些程序（包括模仿者）的一些介绍，参见Hofstadter D.，*Fluid Concepts and Creative Analogies.* New York：Basic Books，1995。

[188] “意义屏障”：Rota, G.-C.，In memoriam of Stan Ulam-The barrier of meaning.*Physica* D，2（1-3），1986，pp.1—3。

[189] “根据笛卡儿的思想，这些基本要素是所有物理模型必要而且充分的组成部分”：Garber, D.，Descartes, mechanics, and the mechanical philosophy.*Midwest Studies in Philosophy* ，26（1），2002，pp.185—204。

[190] “猜想地球就像海绵一样”：Kubrin, D.，Newton and the cyclical cosmos：Providence and the mechanical philosophy.*Journal of the History of Ideas*，28（3），1967。

[191] “直觉泵”：Dennett, D.R.，*Elbow Room*：*The Varieties of Free Will Worth Wanting.* Cambridge, MA：MIT Press，1984，p.12。

[192] “弗勒德和德雷希尔发明了囚徒困境”：对囚徒困境以及博弈论的历史和应用有趣而富有启发的讨论，参见Poundstone, W.，*Prisoner’s Dilemma* .New York：Doubleday，1992。

[193] “每个人都追求自利，使得所有人的利益都受损”：Axelrod, R.，*The Evolution of Cooperation.* New York：Basic Books，1984，p.7。

[194] “公地悲剧”：Hardin, G.，The tragedy of the commons.Science，162，1968，pp.1243—1248.

[195] “合作要如何才能出现？”：Axelrod, R.，*The Evolution of Cooperation.* New York：Basic Books，1984，p.3。

[196] “霍布斯……认为合作只有在存在中央权威的情况下才有可能产生”：霍布斯关于集权政府的论述参见Hobbes, T.，*Leviathan* .1651年首次出版；1991年版由R.Tuck编辑。Cambridge University Press，1991。

[197] “爱因斯坦也提出了类似的观点”：爱因斯坦关于世界政府等问题的思想参见他的著作集，Einstein, A.，*Out of My Later Years* .1950年首次出版；2005年修订版，Castle Books。

[198] “阿克塞尔罗德进行了添加社会规范的囚徒困境实验”：Axelrod, R.，An evolutionary approach to norms.*American Political Science Review*，80（4），1986，pp.1095—1111。

[199] “元规范能促进并保持群体中的合作”：Axelrod, R.，An evolutionary approach to norms.*American Political Science Review*，80（4），1986，pp.1095—1111。

[200] “诺瓦克同他的博士后导师梅一起进行了计算机仿真”：Nowak, M.A.&May, R.M.，Evolutionary games and spatial chaos.*Nature*，359（6398），1992，pp.826—829。

[201] “我们认为，群体中确定的空间结构对合作的进化可能很重要”：Nowak, M.A.&May, R.M.，Evolutionary games and spatial chaos.#Nature#，359（6398），1992，pp.826—829。

[202] “混沌性变化”：Nowak, M.A.&May, R.M.，Evolutionary games and spatial chaos.#Nature#，359（6398），1992，pp.826—829。

[203] “地域性有利于合作”：Sigmund, K.，On prisoners and cells，#Nature#，359（6398），1992，p.774。

[204] “霍兰德曾把这种模型比作‘飞行模拟器’”：Holland, J.H.，*Emergence*：*From Chaos to Order* .Perseus Books，1998，p.243。”

[205] “被用来帮助改进P2P网络”：例如，Hales, D.&Arteconi, S.，SLACER：A Self-Organizing Protocol for Coordination in Peer-to-Peer Networks.*IEEE Intelligent Systems*，21（2），2006，pp.29—35。

[206] “在电子商务中防止欺诈”：例如，参见Kollock, P.，The production of trust in online markets.收录在E.J.Lawler，M.Macy，S.Thyne，&H.A.Walker（编辑），*Advancesin Group Processes*，Vol.16.Greenwich, CT：JAI Press，1999。

[207] “诺瓦克在题为《合作进化的5条规则》的文章中所做的工作”：Nowak, M.A.，Five rules for the evolution of cooperation.*Science*，314（5805），2006，pp.1560—1563。

[208] “新能源金融……最近公布了一份报告”：Liebreich, M.，How to Save the Planet：Be Nice,

Retaliatory, Forgiving，&Clear.White Paper, New Energy Finance, Ltd.，2007。（http：//www.newenergyfinance.com/docs/Press/NEF_WP_Carbon-Game-Theory_05.pdf）

[209] “对政府制定应对恐怖主义、军备控制和环境治理等问题的政策也有影响”：例如，参见Cupitt, R.T.，Target rogue behavior, not rogue states.*The Nonproliferation Review*，3，1996，pp.46—54；Cupitt, R.T.&Grillot, S.R.，COCOM is dead, long live COCOM：Persistence and change in multilateral security institutions.*British Journalof Political Science* 27，7，pp.361—389；以及Friedheim, R.L.，Ocean governance at the millennium：Where we have been, where we should go：Cooperation and discord in the world economy.*Ocean and Coastal Management*，42（9），1999，pp.747—765。

[210] “生物多样性的维持和细菌制造新抗生素的效率问题”：例如，Nowak, M.A.&Sigmund, K.，Biodiversity：Bacterial game dynamics.*Nature* ，418，2002，pp.138—139；Wiener, P.，Antibiotic production in a spatially structured environment.Ecology Letters，3（2），2000，pp.122—130。

[211] “所有模型都是错的，但是有一些有用”：Box, G.E.P.&Draper, N.R.，*Empirical Model Building and Response Surfaces.* New York：Wiley 1997，p.424。

[212] “可重复性是科学积累的基石”：Axelrod R.，Advancing the art of simulation in the social sciences.收录在Conte, R.，Hegselmann, R.，Terna, P.（编辑），*Simulating Social Phenomena.*（Lecture Notes in Economics and Mathematical Systems 456）.Berlin：Springer-Verlag，1997。

[213] “胡伯尔曼和格兰斯重复了诺瓦克和梅的空间囚徒困境模型”：Huberman, B.A.&Glance, N.S.，Evolutionary games and computer simulations.*Proceedings of the National Academy of Science*，USA，90，1993，pp.7716—7718。

[214] “穆科吉、拉詹和斯莱格勒也独立得到了类似的结果”：Mukherji, A.，Rajan, V.，&Slagle, J.R.，Robustness of cooperation.*Nature*，379，1996，pp.125—126。

[215] “诺瓦克、梅和他们的合作者博恩霍艾弗则回应说”：Nowak, M.A.，Bonhoeffer, S.，&May, R.M.，Spatial games and the maintenance of cooperation.Proceedings of the National Academy of Sciences, USA，91，1994，pp.4877—4881；Nowak, M.A.，Bonhoeffer, S.，&May, R.M.，Reply to Mukherji et al.*Nature*，379，1996，p.126。

[216] “加兰和利兹奎尔多发表了他们重复阿克塞尔罗德的规范和元规范模型实验的结果”：Galan, J.M.&Izquierdo, L.R.，Appearances can be deceiving：Lessons learned re-implementing Axelrod’s‘Evolutionary Approaches to Norms.’*Journal of Artificial Societies and Social Simulation*，8（3），2005。（http：//jasss.soc.surrey.ac.uk/8/3/2.html）

[217] “建模的艺术就是去除实在中与问题无关的部分”：Anderson, Nobel Prize acceptance speech，1977。

[218] “在爱希莉亚”：引自Calvino, I.Invisible Cities.New York：Harcourt Brace Jovanovich，1974，p.76。（英文翻译W.Weaver）

[219] “网络科学”：本章部分内容源自Mitchell, M.，Complex systems：Network thinking.*Artificial Intelligence*，170（18），2006，pp.1194—1212.

[220] “米尔格兰姆记述了一个例子”：引自Milgram, S.，The small-world problem.*Psychology Today*，1，1967，pp.61—67。

[221] “后来心理学家柯兰菲尔德研究发现”：参见Kleinfeld, Could it be a big world after all?*Society*，39，2002。

[222] “我们文化的传奇”：Kleinfeld, J.S.，Six degrees：Urban myth?*Psychology Today*，74，March/April 2002。

[223] “当人们发现出人意料的社会关系时”：Kleinfeld, J.S.，Could it be a big world after all? The“six degrees of separation”myth.Society，39，2002。

[224] “‘网络新科学’”：例如，Barabási, A.-L.，Linked：*The New Science of Networks.* Cambridge, MA：Perseus，2002。

[225] “20世纪90年代末的两篇重要文章”：Watts, D.J.&Strogatz, S.H.，Collective dynamics of‘small world’networks.*Nature*，393，1998，pp.440—442；Barabási, A.-L.&Albert, R.，Emergence of scaling in random networks，*Science*，286，1999，pp.509-512。

[226] “一大群饥肠辘辘的物理学家”：Watts, D.J.，*Six Degrees*：*The Science of a Connected Age.* New York：W.W.Norton&Co，2003，p.32。

[227] “网络思维将渗透到人类活动和人类思想的一切领域”：Barabási, A.—L.Linked：*The New Science of Networks.* Cambridge, MA：Perseus，2002，p.222。

[228] “123个链接指向我自己的网站”：本章用到的所有网站入链接数据都来自网页http：//www.microsoft-watch.org/cgi-bin/ranking.html。这个数据只包括来自网站外部的入链接。

[229] “从数学上定义了小世界网络的概念”：参见Watts, D.J.&Strogatz, S.H.，Collective dynamics of‘small world’networks.*Nature*，393，1998，pp.440—442。

[230] “结果图15.8中的规则网络的平均路径长度为15”：这个值是用公式l=N/2 k计算得出。其中l是平均路径长度，N是节点数量，k是各节点的连接度（这里是2）。参见Albert, R.&Barabási, A-L.，Statistical mechanics of complex networks.*Reviews of Modern Physics*，74，2002，pp.48—97。

[231] “平均路径长度一下就降到了9左右”：这个值是根据以下文献给出的结果估计，Newman,

M.E.J.，Moore, C.，&Watts, D.J.，Mean-field solution of the small-world network model.*Physical Review Letters*，84，1999，pp.3201—3204。

[232] “只需很少的随机连接就能产生很大的效应”：Watts, D.J.，Six Degrees：*The Science of a Connected Age.* New York：W.W.Norton，2003，p.89。

[233] “小世界性”：小世界性的正式定义是这样，即使长程连接相对较少，在平均连接度不变的情况下，两个节点之间的最短路径长度（跨越的边的数量）会随网络大小n呈对数增长或更低。

[234] “凯文·贝肯游戏”：参见，例如，http：//en.wikipedia.org/wiki/Six_Degrees_of_Kevin_Bacon。2.“神经学家已经绘制出了这种低等生物的所有神经元和连接”：更多信息参见Achacoso, T.B.&Yamamoto, W.S.，*AY’s Neuroanatomy of C.Elegans for Computation.* Boca Raton, FL：CRC Press，1991。

[235] “产生的网络与许多真实世界网络的度分布并不一样”：瓦特—斯托加茨模型产生的网络连接度呈指数分布，而大多数真实世界中的网络连接度是呈幂律分布。细节参见Albert, R.&Barabási, A—L.，（2002）.Statistical mechanics of complex networks.*Reviewsof Modern Physics*，74，2002，pp.48—97。

[236] “报道2008年苹果价格的网页”：http：//www.americanfruitgrower.com/e_notes/page.php?page=news。

[237] “关于塔斯马尼亚大苹果赛信息的网页”：http：//www.huonfranklincottage.com.au/events.html。

[238] “这个规则实际上只适用于”：网页的入度分布基本符合幂律$k^{-2.3}$，下限kmin=3684[参见Clauset, A.，Shalizi, C.R.，&Newman, M.E.J.，Power-law distributions in empirical data.Preprint，2007（http：//arxiv.org/abs/0706.1062）。]为了便于讨论，这一章中将式子简化近似为k^{-2}；$k^{-2.3}$分布的图与这一章给出的图非常相似。

[239] “神经学家又绘制出了猫、恒河猴等动物甚至人类的一些高级功能脑区的连接结构”：例如，参见Bassett, D.S.&Bullmore, D.，Small-world brain networks.*The Neuroscientist*，12，2006，pp.512—523；以及Stam, C.J.&Reijneveld, J.C.，Graph theoretical analysis of complex networks in the brain.*Nonlinear Biomedical Physics*，1（1），2007，p.3。

[240] “基因调控网络”：对于网络思想应用于基因调控的更多细节，参见Barabási, A.-L.&Oltvai, Z.N.，Network biology：Understanding the cell’s functional organization.*Nature Reviews*：*Genetics*，5，2004，pp.101—113。

[241] “巴拉巴西和他的同事仔细研究了43种生物的代谢网络结构”：Jeong, H.，Tombor, B.，Albert, R.，Oltvai, Z.N.，&Barabási, A.-L.，The large scale organization of metabolic

networks.Nature，407，2000，pp.651—654。其他研究代谢网络结构的例子包括Fell, D.A.&Wagner, A.，The small world of metabolism.*Nature Biotechnology*，18，2000，pp.1121—1122；以及Burgard, A.P.，Nikolaev, E.V.，Schilling, C.H.，&Maranas, C.D.，Flux coupling analysis of genome-scale metabolic network reconstructions.*Genome Research*，14，2004，pp.301—312。

[242] “虽然后来的研究否认了杜格斯是北美的传染源”：参见，例如，Robbins, K.E.，Lemey, P.，Pybus, O.G.，Jaffe, H.W.，Youngpairoj, A.S.，Brown, T.M.，Salemi, M.，Vandamme, A.M.，&Kalish, M.L.，U.S.human immunodeficiency virus type 1 epidemic：Date of origin, population history, and characterization of early strains.*Journalof Virology*，77（11），2003，pp.6359—6366。2.“最近，一个由社会学家和物理学家组成的团队分析了瑞典性行为调查数据”：Liljeros, F.，Edling, C.R.，Nunes Amaral, L.A.，Stanely, H.E.，&Aberg, Y.，The web of human sexual contacts.Nature，441，2001，pp.907—908。

[243] “其他对性关系网络的研究也得出了类似结论”：例如，Schneeberger, A.，Mercer, C.H.，Gregson, S.A.，Ferguson, N.M.，Nyamukapa, C.A.，Anderson, R.M.，Johnson, A.M.，&Garnett, G.P.，Scale-free networks and sexually transmitted diseases：A description of observed patterns of sexual contacts in Britain and Zimbabwe.*Sexually Transmitted Diseases*，31（6），2004，pp.380—387。

[244] “另一个网络科学家团体提出了一个巧妙而简单的方法”：Cohen, R.，ben-Avraham, D.，&Havlin, S.，Efficient immunization strategies for computer networks and populations.*Physics Review Letters*，91（24），2003，p.247—901。

[245] “杀毒应当重点针对邮件通信录很长的用户的计算机”：Newman, M.E.J.，Forrest, S.，&Balthrop, J.，Email networks and the spread of computer viruses.*Physical Review E*，66，2002，p.035—101。

[246] “生物学界最近对这个问题有很多争议”：参见，例如，Montoya, J.M.&Solé，R.V.，Small world patterns in food webs.*Journalof Theoretical Biology*，214（3），2002，pp.405—412；Dunne, J.A.，Williams, R.J.，&Martinez, N.D.，Food-web structure and network theory：The role of connectance and size.*Proceedings of the National Academy of Science*，USA，99（20），2002，pp.12917—12922；以及Dunne, J.A.，The network structure of food webs.收录在M.Pascual&J.A.Dunne（编辑），*Ecological Networks*：*Linking Structure to Dynamics in Food Webs*. New York：Oxford University Press，2006，pp.27—86。

[247] “无尺度网络是如何产生的”：这一节部分来自Mitchell, M.，Complex systems：Network thinking.*Artificial Intelligence*，170（18），2006，pp.1194—1212。

[248] “巴拉巴西和艾伯特提出了一种网络生长机制”：Barabási, A.—L.&Albert, R.，Emergence of scaling in random networks, Science，286，1999，pp.509—512。

[249] “这种机制以及所产生的幂律在之前至少被独立发现过三次”：Yule, G.U.，A mathematical theory of evolution, based on the conclusions of Dr.J.C.Willis.*Philosophical Transactionsofthe Royal Society of London* ，Ser.B 213，1924，pp.21—87；Simon, H.A.，On a class of skew distribution functions.*Biometrika* ，42（3—4），1955，p.425；以及Price, D.J.，Networks of scientific papers.*Science* ，149，1965，pp.510—515。

[250] “科技文献引用网络的增长”：例如，参见Redner, S.，How popular is your paper?An empirical study of the citation distribution.*European Physical Journal B* ，4（2），1998，pp.131—134。2.“作家格拉德威尔所说的引爆点”：Gladwell, M.，*The Tipping Point*：*How Little Things Can Make a Big Difference.* Boston：Little, Brown，2000。

[251] “用这种方法确定为‘无尺度’的一些网络后来又被发现其实不是无尺度分布”：Clauset, A.，Shalizi, C.R.，&Newman, M.E.J.，Power-law distributions in empirical data.Preprint，2007（http：//arxiv.org/abs/0706.1062）。

[252] “现在幂律的普遍性可能被高估了”：Keller, E.F.，Revisiting‘scale-free’networks.*Bio Essays* ，27，2005，pp.1060—1068。

[253] “我们对幂律的迷恋是一种耻辱”：Shalizi, C.，Networks and Netwars，2005。论文位于http：//www.cscs.umich.edu/～crshalizi/weblog/347.html。

[254] “产生幂律的方法有许多，每种都有道理”：Shalizi, C.，Power Law Distributions，1/f noise, Long-Memory Time Series，2007。论文位于http：//cscs.umich.edu/～crshalizi/notebooks/power-laws.html。

[255] “几乎每隔一天都有讲座介绍产生幂律分布的新机制”：对这些机制的综述，参见Mitzenmacher, M.，A brief history of generative models for power law and lognormal distributions.*Internet Mathematics* ，1（2），2003，pp.226—251；以及Newman, M.E.J.，Power laws, Pareto distributions and Zipf’s law.*Contemporary Physics* ，46，2005，pp.323—351。

[256] “据报道，由于天气过于炎热”：这次的连锁失效及其原因的详细报道参见U.S.-Canada Power System Outage Task Force’s Final Report on the August 14，2003 Blackout in the United States and Canada：Causes and Recommendations（https：//reports.energy.Gov/）。

[257] “美国海关计算机系统崩溃了近10个小时”：参见Schlossberg, D.“LAX Computer Crash Strands International Passengers.”Consumer Affairs.com，2007年8月13日（http：//www.consumeraffairs.com/news 04/2007/08/lax_computers.html）；以及Schwartz, J.，“Who Needs Hackers？”New York Times，2007年9月12日。

[258] “美国长期资本管理公司”：参见，例如，Government Accounting Office，*Long-Term Capital Management*：*Regulators Need to Focus Greater Attention on Systemic Risk.* Report to Congressional Request，1999（http：//www.gao.gov/cgi-bin/getrpt?GGD-00-3）；以及Coy, P.，

Woolley, S.，Spiro, L.N.，&Glasgall, W.，Failed wizards of Wall Street.Business Week，1998年9月21日。

[259] “威胁来自复杂性本身”：安东诺普洛斯，引自Schwartz, J.，“Who Needs Hackers？”*New York Times*，2007年9月12日。

[260] “自组织临界性”：对SOC的介绍，参见Bak, P.，*How Nature Works*：*The Science of Self-Organized Criticality* .New York：Springer，1996。

[261] “高容错性”：对HOT的介绍，参见Carlson, J.M.&Doyle, J.，Complexity and robustness.*Proceedingsofthe National Academyof Science*，USA99，2002，pp.2538—2545。

[262] “对于网络的动力学之谜”：Watts, D.J.，*Six Degrees*：*The Science of a Connected Age*. New York：W.W.Norton，2003，p.161。

[263] “表面积与体积的2/3次幂呈比例”：记体积为V，表面积为S，半径为r。V正比r^3 ，因此V的立方根正比于半径。表面积正比于r^2，因此正比于体积立方根的平方，也就是$V^{2/3}$ 。

[264] “在双对数图上幂律曲线表现为直线，而直线的斜率则等于幂律的指数”：在这个例子中，幂律为：代谢率∝体重$^{3/4}$ ，两边同时取对数，得到：log（代谢率）∝3/4 log（体重）。这是斜率为3/4的直线方程，如果以log（代谢率）和log（体重）为轴作图，画出来就是图17.2那样。

[265] “恩奎斯特后来用‘放烟火’来描述他们的数学结果”：Grant, B.，*The powers that be.The Scientist*，21（3），2007。

[266] “你应当在两个不同的尺度上思考”：韦斯特，引自Mackenzie, D.，Biophysics：New clues to why size equals destiny.*Science*，284（5420），1999，pp.1607—1609。

[267] “推导出指数为3/4的模型细节相当复杂”：对代谢比例模型的一个专业但不难理解的阐释参见West, G.B.&Brown, J.H.，Life’s universal scaling laws.*Physics Today*，57（9），2004，pp.36—43。

[268] “虽然生物是三维的”：West, G.B.，Brown, J.H.，&Enquist, B.J.，The fourth dimension of life：Fractal geometry and allometric scaling of organisms.*Science*，284，pp.1677—1679。

[269] “有统一整个生物学的潜力”：Grant, B.，The powers that be.*The Scientist*，21（3），2007。

[270] “对于生物学的重要性就好比牛顿的发现对于物理学的重要性”：Niklas, K.J.，Size matters！*Trends in Ecology and Evolution*，16（8），2001，p.468。

[271] “可以预见，广义代谢理论的涌现对于生物学的重要性将类似于遗传理论”：West,

G.B.&Brown, J.H.，The origin of allometric scaling laws in biology from genomes to ecosystems：Towards a quantitative unifying theory of biological structure and organization.*Journalof Experimental Biology*，208，2005，pp.1575—1592。

[272] “甚至有人认为克莱伯根本就是错的”：对代谢比例理论的各种批评的综述，参见Agutter P.S&Wheatley, D.N.，Metabolic scaling：Consensus or Controversy?*Theoretical Biology and Medical Modeling*，18，2004，pp.283—289。

[273] “人们对于所涉及的生理细节了解得越多”：H.Horn，引自Whitfield, J.，All creatures great and small.*Nature*，413，2001，pp.342—344。

[274] “事情简单当然很好”：穆勒—兰道，引自Grant, B.，The powers that be.*The Scientist*，21（3），2007。

[275] “另外还有人认为代谢比例理论的数学有错误”：例如，Kozlowski, J.&Konarzweski, M.，Is West, Brown and Enquist’s model of allometric scaling mathematically correct and biologically relevant?*Functional Ecology*，18，2004，pp.283—289。

[276] “研究组坚持自己的立场，他们对吹毛求疵的批评意见感到沮丧”：例如，参见West, G.B.，Brown, J.H.，&Enquist, B.J.，Yes, West, Brown and Enquist’s model of allometric scaling is both mathematically correct and biologically relevant.（Reply to Kozlowski and Konarzweski，2004.）*Functional Ecology*，19，2005，pp.735—738；以及Borrell, B.，Metabolic theory spat heats up.*The Scientist*（News），November 8，2007。（http：//www.the-scientist.com/news/display/53846/）。

[277] “我的内心不会向这些在我脚边乱吠的小狗屈服”：韦斯特，引自Grant, B.，The powers that be.*The Scientist*，21（3），2007。

[278] “我想韦斯特和恩奎斯特等人不会一直重复他们的观点”：穆勒—兰道，引自Borrell, B.，Metabolic theory spat heats up.*The Scientist*（News），November 8，2007。（http：//www.the-scientist.com/news/display/53846/）

[279] “比‘正态’还要正态”；“在复杂的自然和工程系统中获得的数据中发现（幂律）分布”：Willinger, W.，Alderson, D.，Doyle, J.C.，&Li, L.，More‘normal’than normal：Scaling distributions and complex systems.收录在R.G.Ingalls等，*Proceedingsofthe2004Winter Simulation Conference*，pp.130—141.Piscataway, NJ：IEEE Press，2004。

[280] “这个关系现在被称为齐普夫定律”：齐普夫最初发表这项研究是在一本书中：Zipf, G.K.，*Selected Studies of the Principle of Relative Frequency in Language.* Cambridge, MA：Harvard University Press，1932。

[281] “曼德布罗特从信息量的角度提出了不同的解释”：Mandelbrot.B.，An informational theory

of the statistical structure of languages.收录在W.Jackson（编辑），*Communicaiton Theory* ，Woburn, MA：Butterworth，1953，pp.486—502。

[282] “西蒙也提出了一种解释”：Simon, H.A.，On a class of skew distribution functions.*Biometrika* ，42（3—4），1955，p.425。

[283] “对于曼德布罗特和西蒙的解释哪个正确，争论很激烈”：Mitzenmacher, M.，A brief history of generative models for power law and lognormal distributions.*Internet Mathematics* ，1（2），2003，pp.226—251。

[284] “心理学家乔治·米勒使用简单的概率论证明”：Miller, G.A.，Some effects of intermittent silence.*The American Journal of Psychology* ，70，1957，pp.311—314。

[285] “可动遗传因子”：关于可动遗传因子对大脑多样性的可能作用，有一篇专业文献是Muotri, A.R.，Chu, V.T.，Marchetto, M.C.N.，Deng, W.，Moran, J.V.&Gage, F.H.，Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.*Nature* ，435，2005，pp.903—910。

[286] “最近一组科学哲学家和生物学家进行了一项调查”：Pearson, H.，What is a gene?*Nature* ，441，2006，pp.399—401。

[287] “对分子遗传学越专长的学者”：同上。

[288] “胚胎发育期缺乏必需的甲基化”：参见，例如，Dean, W.，Santos, F.，Stojkovic, M.，Zakhartchenko, V.，Walter, J.，Wolf, E.，&Reik, W.，Conservation of methylation reprogramming in mammalian development：Aberrant reprogramming in cloned embryos.*Proceedingsofthe National Academy of Science* ，USA，98（24），2001，pp.13734—13738。

[289] “DNA转录为RNA之后很大部分最终都没有被译码成蛋白质”：参见，例如，Mattick, J.S.，RNA regulation：A new genetics?Nature Reviews：Genetics，5，2004，pp.316—323；Grosshans, H.&Filipowicz, W.，The expanding world of small RNAs.*Nature* ，451，2008，pp.414—416。

[290] “非编码RNA的作用是目前遗传学中一个非常活跃的研究领域”：对这个领域目前的一些研究和争议的讨论，参见Hüttenhofer, A.，Scattner, P.，&Polacek, N.，Non-coding RNAs：Hope or Hype?*Trends in Genetics* ，21（5），2005，pp.289—297。

[291] “基因独立运作的想法是1976年之后形成的”：Caruso, D.，A challenge to gene theory, a tougher look at biotech.*New York Times* ，2007年7月1日。

[292] “编码了特定的功能性产物”：美国专利权法，引自同上。

[293] “基因组网络化的证据实际上毁掉了对当今生物技术产品商业化进行的所有官方风险评估的科学基础”：Caruso, D.，A challenge to gene theory, a tougher look at biotech.*New York Times* ，

2007年7月1日。

[294] “我们的DNA超过90%与老鼠一样”：在发表的文献中对人类与其他物种DNA重叠部分的估计各不相同。不过，“超过90%一样”应该是相当可靠的估计。

[295] “‘黑箱’，在其中可以将遗传信息变成三维的、活的动物”：Carroll, S.B.，*Endless Forms Most Beautiful*：*The New Science of Evo Devo and the Making of the Animal Kingdom.* New York：W.W.Norton，2005，p.7。

[296] “讽刺的是……一直被视为垃圾的[DNA]却藏有人类复杂性的秘密”：参见视频Gene Regulation；Science，319，no.5871，2008。

[297] “为了证明BMP4基因确实能触发生长粗壮、能打开坚果的喙”：引自Yoon, C.K.，From a few genes, life’s myriad shapes.*New York Times*，2007年6月26日。

[298] “格林做了一个古怪而富有启发的实验”：这项研究参见Travis, J.，Eye opening gene.*Science News Online*，1997年3月10日。

[299] “这个结论在进化生物学家中仍然极具争议”：对这个争议的讨论，参见，例如，Erwin, D.H.，The developmental origins of animal bodyplans.收录在S.Xiao&A.J.Kaufman（编辑），*Neoproterozoic Geobiologyand Paleobiology*，New York：Springer，2006，pp.159—197。

[300] “世界级的知识即兴演奏家”：Horgan, J.，From complexity to perplexity.*Scientific American*，272，1995年6月，pp.74—79。

[301] “控制从受精卵到成人的发育的基因组网络”：Kauffman, S.A.，*At Home in the Universe*.New York：Oxford University Press，1995，p.26。

[302] “生命存在于混沌的边缘”：同上，p.26。

[303] “继承了达尔文传统的大部分生物学家都认为”：Kauffman, S.A.，*At Home in the Universe*.New York：Oxford University Press，1995，p.26，p.25。

[304] 译注：美国漫画家鲁布·戈德堡（Rube Goldberg）在他的作品中创作出各种复杂的机械完成各种简单的任务，后来人们以“鲁布·戈德堡机器”命名这种装置，意指“以极为繁复而迂回的方法去完成实际上或看起来可以很容易做到的事情”，也指“极为混乱或复杂的系统”。

[305] “候选的热力学第四定律”：Kauffman, S.J.，*Investigations*.New York：Oxford University Press，2002，p.3。

[306] “考夫曼的《秩序的起源》一书”：Kauffman, S.A.，*The Origins of Order*.New York：Oxford University Press，1993。

[307] “对选择在进化论中的地位进行根本性的重新解读”：Burian, R.M.&Richardson, R.C.，Form and order in evolutionary biology：Stuart Kauffman’s transformation of theoretical biology.收录在*PSA*：*Proceedings of the Biennial Meeting of the Philosophy of Science Association*，Vol.2：Symposia and Invited Papers，1990，pp.267—287。

[308] “他的方法开启了新的未来”：同上。

[309] “对整个生物学进行建模的第一次严肃尝试”：Bak, P.，*How Nature Works*：*The Science of Self-Organized Criticality*. New York：Springer，1996。

[310] “危险的诱惑力”：Dover, G.A.，On the edge.*Nature*，365，1993，pp.704—706。5.“有时候就好像在虚幻的空间悠然漫步”：同上。

[311] “噪声对RBN的行为有明显的影响”：例如，参见Goodrich, C.S.&Matache, M.T.，The stabilizing effect of noise on the dynamics of a Boolean network.Physica A，379（1），2007，pp.334—356。

[312] “加入捍卫达尔文主义的战斗现在已经成为进化生物学家必需的社会责任”：Hoelzer, G.A.，Smith, E.，&Pepper, J.W.，On the logical relationship between natural selection and self-organization.*Journal of Evolutionary Biology*，19（6），2007，pp.1785—1794。

[313] “进化生物学家麦舒教了我一个思考这些问题的好办法”：麦舒，私人通信。

[314] “进化生物学处于知识混沌状态中”：麦舒，私人通信。

[315] “我要将混沌放入十四行诗”：收录在Millay, E.St.Vincent，*Mine the Harvest*：*A Collection of New Poems.New York*：Harper，1949，p.130。

[316] “霍根在世界顶级科普杂志《科学美国人》上发表了一篇文章”：Horgan, J.，From complexity to perplexity.*Scientific American*，272，1995年6月，pp.74—79。

[317] “《科学的终结》”：Horgan, J.，*The End of Science*：*Facing the Limits of Knowledge in the Twilight of the Scientific Age*. Reading, MA：Addison-Wesley，1996。

[318] “希望物理学能彻底理解基本力从而完结的想法是没有根基的”：Crutchfield, J.P.，Farmer, J.D.，Packard, N.H.，&Shaw, R.S.，Chaos.*Scientific American*，255，1986年12月。

[319] “人们相爱不能怪万有引力”：这句话被普遍认为是爱因斯坦说的。不过，这显然是改自他的原话：“堕入情网并不是人类所做的最愚蠢的事——但万有引力规则不需为此负责。”引自Dukas, H.&Hoffmann B.（编辑），*Albert Einstein, The Human Side*：*New Glimpsesfrom His Archives*. Princeton, NJ：Princeton University Press，1979，p.56。

[320] “继控制隐喻之后”：Gordon, D.M.，Control without hierarchy.*Nature*，446（7132），

2007，p.143。

[321] “一门新的学科，控制论”：关于控制论的迷人历史可以参考Aspray, W.，*John von Neumann and the Origins of Modern Computing.* Cambridge, MA：MIT Press，1990；以及Heims, S.*The Cybernetics Group.* Cambridge, MIT Press，1991。

[322] “整个控制和通信的理论”：Wiener, N.*Cybernetics* .Cambridge, MA：MIT Press，1948，p.11。

[323] “阿什比的《大脑设计》”：Ashby, R.H.，*Design for a Brain* .New York：Wiley，1954。

[324] “麦卡洛克和皮茨提出了神经元模型作为逻辑器件”：Mc Culloch, W.&Pitts, W.，A logical calculus of ideas immanent in nervous activity.*Bulletinof Mathematical Biophysics* 5，1942，pp.115—133。

[325] “米德和贝特森将控制论的思想应用到心理学和人类学”：参见，例如，Bateson, G.，*Mind and Nature*：*A Necessary Unity.* Cresskill, NJ：Hampton Press，1979。

[326] “维纳的著作《控制论》和《人有人的用处》”：Wiener, N.*Cybernetics*：*Or the Control and Communication in the Animal and the Machine.* Cambridge, MA：MIT Press，1948；Wiener, N.*The Human Use of Human Beings.* Boston：Houghton Mifflin，1950。

[327] “我一生经历的最重要的两次历史事件是凡尔赛条约的签订和控制论的发现”：贝特森，引自Heims, S.，*The Cybernetics Group.* Cambridge：MIT Press，1991，p.96。

[328] “极为空洞无物”：德尔布吕克，引自Heims, S.，*The Cybernetics Group* .Cambridge：MIT Press，1991，p.95。

[329] “非常精英的团体在一起闲谈”：萨维奇，引自Heims, S.，*The Cybernetics Group* ，1991，p.96。

[330] “最后维纳统一控制和通信科学的愿望没有实现”：Aspray, W.，*John von Neumann and the Origins of Modern Computing.* Cambridge：MIT Press，1990，pp.209—210。

[331] “一般系统论”：参见Von Bertalanffy, L.，*General System Theory*：*Foundations, Development, Applications* ，New York：G.Braziller，1969；或Rapoport, A.*General System Theory*：*Essential Concepts and Applications* ，Cambridge, MA：Abacus Press，1986。

[332] “对一般性‘系统’有效的原则进行形式化和演绎”：Von Bertanlanffy, L.，An outline of general system theory.*The British Journal for the Philosophy of Science* ，1（92），1950，pp.134—165。

[333] “生物学家马图拉纳和维埃拉试图用自创生的概念统一前两条主线”：参见，例如，

Maturana, H.R.&Varela, F.J.，*Autopoiesis and Cognition*：*The Realization of the Living.Boston*，MA：D.Reidel Publishing Co.，1980。

[334] “哈肯的协同学和普里高津的耗散结构和非平衡系统理论”：参见Haken, H.，*The Science of Structure*：*Synergetics*，New York：Van Nostrand Reinhold，1984；以及Prigogine, I.*From Being to Becoming*：*Time and Complexity in the Physical Sciences*，San Francisco：W.H.Freeman，1980。

[335] “复杂性的词汇表”，“涉及在各种现象中反复遇到的机制的一系列概念”：Nicolis, G.&Prigogine, I.，*Exploring Complexity*，New York：W.H.Freeman and Co.，1989，p.x。

[336] “我认为我们可能缺乏与微积分相当的新概念体系”：Strogatz, S.，Sync：*How Order Emerges from Chaos in the Universe, Nature, and Daily Life.* New York：Hyperion，2004，p.287。

[337] 他受困于语言的混乱”：Gleick, J.，*Isaac Newton*，New York：Pantheon Books，2003，pp.58—59。

[338] “物理学家巴克在动力系统理论和相变概念的基础上提出了自组织临界性的概念”：参见Bak, P.，*How Nature Works*：*The Science of Self-Organized Criticality*. New York：Copernicus，1996。

[339] “物理学家克鲁奇菲尔德提出了计算力学”：参见，例如，Crutchfield, J.P.，The calculi of emergence.*Physica D*，75，1994，11—54。

[340] “复杂性背后的简单性”：“我不会在乎复杂性前面的简单性，但我会为复杂性背后的简单性奉献一生。”这句名言通常被认为是霍姆斯（Oliver Wendell Holmes）说的，但我没有在他的著作中找到出处。也有人认为这句话是诗人霍普金斯（Gerald Manley Hopkins）说的。

[341] “不敢远离海岸线，就别想发现新大陆”：Gide, A.，*The Counterfeiters* .D.Bussy英文翻译.New York：Vintage，1973，p.353.原版：Journal des Faux-Monnayeurs.Paris：Librairie Gallimard，1927。

[342] 译注：中文版为《链接：网络新科学》，湖南科学技术出版社。

第一推动丛书: 综合系列

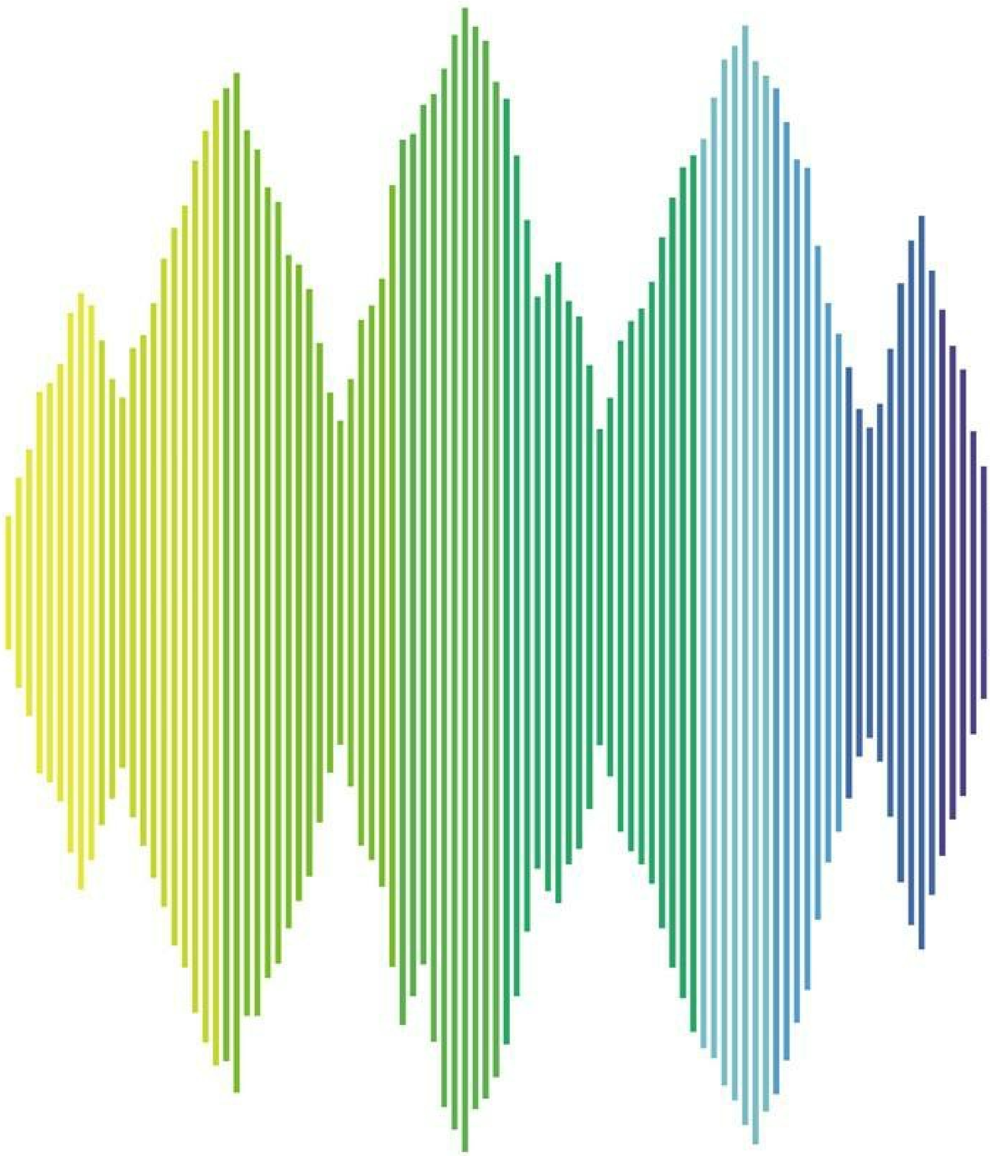
复杂的引擎

[美] 约翰 · E.梅菲尔德 著
唐璐 译

The Polytechnique Series

The Engine of Complexity

John E. Mayfield



湖南科学技术出版社

版权信息

复杂的引擎

著者：[美]约翰·E.梅菲尔德

译者：唐璐

责任编辑：吴炜　孙桂均　李蓓

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794611

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
序
引言
第1章 问题
第2章 计算
第3章 免费的结构
第4章 目的性结构
第5章 无概率性和复杂引擎
第6章 算法进化
第7章 身体内的进化
第8章 控制循环
第9章 复杂系统
第10章 人类学习和创造性
第11章 文化的进化
第12章 复杂性的进化
第13章 过去和现在
第14章 未来
致谢
注释
名词解释
参考文献
译名表

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学

启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认"科学是有用的"，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的"第一推动"。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为"第一推动"，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。

虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

序

像我们自身这样复杂的事物是怎么可能出现的？这个问题困扰了数个世纪来最杰出的一些思想家。复杂性的本质问题带来了理论和方法上的挑战。人们对复杂性提出了许多定义，其中很多都不能量化，无法对这种属性进行度量。我认为，信息、复杂和进化的概念在其中相互缠绕，只有综合考虑，对它们的认识才可能形成突破。而将它们结合到一起的是计算的概念。

极为复杂的事物随处可见。动植物就是明显的例子，还有许多人类发明的事物，例如航天飞机、美国国防部，以及现代汽车制造业背后的供应链。一些无生命、非人造的事物显得也很复杂，例如星系和山脉。有没有可行的方法比较人类和银河系的复杂性？或者珊瑚礁和超级计算机的复杂性？

解决与复杂性有关的问题不是件容易的事情。传统方法——常被称为复杂性科学——采取的是高度数学化的方法。许多专家的背景是物理学，虽然经常会涉及经济和社会系统，但研究的对象多来自非生物领域。我选取了不同的路径，一条偏向哲学和生物学，另一条避开数学，两条路径互为补充。非数学化的方法解释了数学方法曾试图解释的许多事物，反过来又说明了传统的数学为何有时候无效。这本书虽然没有什么公式，但并不是因为反感数学。书中的思想建立在计算理论的基础之上，这是高度专业化的数学分支。不包含公式，是因为在理解相关概念时，它们不是必需的，并且数学推导会给许多读者造成障碍。我希望能让非专业读者也能理解一些基本的计算概念，这些概念虽然本质上极为数学化，但其主要思想却可以用非数学的语言进行阐释。

理查德·道金斯于1986年出版的获奖图书《盲眼钟表匠》部分触及了复杂性问题。他意识到，我们真正想知道的不是如何度量复杂性，而是复杂事物是如何形成的。他指出，一些复杂事物有预先设计，而另一些则没有。计算机的出现是因为工程师的设计。一块嵌满水晶的岩石也很复杂，却没有科学意义上的预先设计；它只是一定条件下一系列自然力量长期作用的产物，而这些条件本身又是一系列长期的自然过程的结

果。岩石和计算机都遵循物理定律，但创造计算机还需要一些东西：精确组织的特定信息。这一点将计算机与非生物、非人造的事物区分开来，从复杂性科学的视角来看，也正是这一点让它不一样。

生物也给人以设计感。然而，水母的构造并非由工程师设计。虽然过去60年，分子遗传学家对水母的结构已经研究得非常清楚，但水母仍然是没有预先设计的。当然现在我们知道了，是水母的DNA编码信息与化学和物理定律的协同运作使得水母的形成成为可能。对于生物，必须有描述其自身并且经常是存于其自身的附加信息，这一点将生命的复杂性与岩石、天气、太阳系这类事物的复杂性清晰地区分开来。因此我们必须追问，这些“附加”的信息从何而来？有意义的信息不会凭空产生，它肯定是来自某个地方。

再次归功于遗传学家，他们研究了达尔文的自然选择过程是如何作用于由不同个体组成的种群，发现生物种群中成员的DNA会随时间变化。个体的DNA不会变化，但种群的DNA会随时间改变。DNA的变化来自突变，然后经过自然选择累积，点滴增加的信息使得水母得以适应环境生存。种群中遗传物质的随机变化（生殖过程中产生的突变）受自然选择的长期作用，就是创造所有生物所需的附加描述信息的来源。但是当工程师设计新的计算机或悬索桥时，附加的描述信息又来自哪里呢？都是工程师凭空想出来的吗？显然不是，后面会解释，当被视为计算时，文化和技术领域遵循的是与生物界同样的运作模式。这意味着让众多复杂事物成为可能的所有附加信息是来自对随机变化的累积选择。我将这个逻辑范式称为“复杂引擎”。同样的引擎驱动了生命的进化以及社会和技术的进步，我猜想人类的学习和创造性也是受其驱动。

信息是个时髦的词：“我们生活在数字（信息）时代”，“生物学正在成为信息科学”，“最有希望统一广义相对论和量子力学的是信息理论”——在这本书中，将许多事物具有的特殊复杂性与岩石、火星或半人马座阿尔法星这类事物的普通复杂性区分开来的是“附加信息”。

根据我们使用手机、计算机的经验，以及我们回应他人话语的基本能力，很显然信息可以被处理、操作和用于达成某种目的。对其进行研究的科学被称为计算机科学。这个学科是众多从不同角度解释世界的学科中的新成员。对计算的数学分析区分了什么是可计算的和不可计算的。一些人认为，从广义的信息和信息处理的角度来看，当我们观察世

界的变化时，我们实际是在观察宇宙的计算。在书中我将阐明，这种相对较新的观点是怎样将进化的思想推广到之前只是在很模糊的意义上被视为具有进化性的领域，以及在许多并不涉及DNA的领域，进化计算是如何运作的。

计算的观点能引申出更广义的进化论，同时也能回答，对于信息没有编码在DNA中的复杂事物，其所需的描述信息来自哪里。答案是随机性：随机事件和随机选择是所有目的性信息的最终来源。这看似疯狂，并且完全有违直觉——这也正是促使我写这本书的动机。

引言

复杂性引出了许多问题。列举少许为例：为什么许多复杂事物看上去都具有目的性或以某种巧妙的方式适应某物？复杂事物与简单事物有何本质区别？复杂事物在理论上是如何可能的？概率上完全不可能的事物到底是如何产生的？对这类问题的疑惑促使我去做研究和写这本书。通过研读前人的成果并不断提升自己的认识，我意识到无法回避背后的进化机制。这里所说的进化，比生物课上讲授的物种起源的过程要更为广义。我说的是许多表面上看似无关的活动背后共同的逻辑。这种逻辑可以视为一种特殊的计算策略，这种策略的执行会产生出原本不可能出现的复杂事物。

许多学者会将这种计算的观点归类于哲学解释的“信息学派”。在我看来，计算的观点让许多原本很“模糊的”概念和讨论变得精确。例如“进化”的精确定义就可以不依赖于生命科学。我给出的定义就可以适用于实现了进化计算的所有系统。用计算来定义进化可以使这个概念更为广义，并且不失精确性。其他好处包括，可以将计算的概念推广到看似非计算的领域，澄清了热力学对进化过程的作用，也澄清了复杂性的研究。

大部分图书馆都藏有关于生物进化和以各种方式解释我们周围的物理世界的优秀书籍。然而虽然有这么好的资源，最近由美国国家科学基金资助的一项调查却发现，只有45%的美国人认为人类是由其他动物进化而来，只有33%的人接受大爆炸是宇宙的起源。如果我们的文化排斥科学技术，这样的结果还好理解，问题是我们不是。美国人早于其他国家很多年就接受了科学发现带来的技术成果，并且经济也因此而受益。讽刺的是，科学在总体上很受推崇，但科学知识的一些部分却由于非科学的原因受到排斥。最为典型的就是对生物尤其是我们自己是源自进化的观点的排斥。从计算的角度认识进化可以凸显出这种排斥的矛盾性，一个人如果承认计算机和大部分人类技术都服从进化机制，就没有理由否认生命进化。书中我将阐明，不仅生物的存在有赖于进化计算，你的身体抵抗感染的能力，人类学习新概念的能力，支撑现代生活的技术，以及作为社会基石的各种组织都有赖于此。

近年来，一个新的科学和数学研究领域对我们的日常生活越来越重要。这个领域就是计算机科学。数字革命正迅速改变人们的行为和思维方式。与此同时，这个新的视角也在改变科学家们对物理、生物学和工程等传统领域的理解和认识，并且帮助我们将它们紧密结合到一起。

将技术、生物学和计算机连接到一起的两个重要概念是信息和进化。我认为，要正确理解进化，就必须借助信息的概念。这是因为进化本质上就是积累信息。让人吃惊的是，信息的本质问题是哲学家、物理学家和生物学家持续不断的争议和混淆的源头。信息一词在不同领域甚至在同一领域中都并不总是意味着相同的事情，虽然这些领域有共同的主题。有趣的是，在计算机科学中似乎没有什么争议。这可能是因为没有清晰的定义这个领域就不可能存在，也有可能是因为计算机科学对实在的本质非常重要。

我是生物学家，同时又对一些更大的问题感兴趣，信息的问题至少有3个方面让我感兴趣。第一，任何一个现代生物学家都认识到，如果不彻底理解DNA中的信息是如何引导生物体的形成，就不可能彻底理解生命。第二，从计算的角度更容易理解进化的过程。根据这种思想，只要遵循特定的信息操作和累积策略，进化就会发生。在书中我将这个策略称为“复杂引擎”。第三，任何有意义的复杂性，无论是不是生物，都只有通过广义的计算过程才有可能实现。这个观点看似不对甚至荒谬，但我希望随着阅读的深入，我能说服你接受这一点。

生命和人类的公司似乎都表现出了高度的创新性。对此进行深入阐释是本书的主题之一。大多数人都会认为，从创新的角度来说，刺猬要比同等大小的花岗岩复杂得多。如果将进化视为计算，就能精确解释为什么刺猬要比花岗岩复杂。

进化计算一旦在某个物理系统中实现，这种“创新计算”往往能产生出原本不可能存在的复杂产物。事实上，进化和创造性的联系极为紧密，很有可能我们遇到的所有具有创造性的复杂事物都只有通过进化计算才有可能。

将达尔文的思想推广到非生物学领域，尤其是社会科学，并不新鲜；但我的一个目标是寻找进化的精确定义，从而回答以下问题：（在达尔文的意义上）过程X到底是不是进化？我在第5章给出的计算性定

义达成了这个目标。从信息的角度理解进化也不新鲜，但即便是我的很多生物学同行，虽然已经是很成功的科学家，并且也完全接受达尔文的观点，仍然有许多人并没有认识到进化计算的威力以及它在生物学以外的作用。

从2500年前的柏拉图和亚里士多德开始，复杂性的起源就一直让许多最杰出的哲学家着迷 [1] 。这方面目前最活跃的领域之一是宇宙学，物理学家在探讨关于宇宙起源和本质的基本问题。目前讨论的一个热点是尝试将已经证明了的理论计算机科学的原理应用到物理学。用MIT的理论物理学家塞思·劳埃德的话说，“我们可以从数学上证明，一个计算的宇宙必然——有高度的可能性——产生越来越复杂的结构。 [2] ”还有观点认为，没有计算就不可能有复杂的结果。这里的计算定义为操作信息的过程。

我在这本书中呈现的思考开始于20年前，1993年我第一次读了道金斯的《盲眼钟表匠》 [3] 。当时我在给大一新生教生物学导论。我采用的教材详细介绍了进化，我讲起来没什么困难，大部分初次接触这个主题的学生也能够接受。但我还是感觉到这种标准的进化阐释似乎缺少了某种重要的东西。

为了丰富教学内容，我读了道金斯的书。坦白说，我自己也想学习一下。我在大学的专业是物理学，研究生专业则是生物物理学。我没有接受过关于进化的科班训练。不过，生物进化的事实不是问题。我的专业主要是研究DNA序列，对于这个领域的人来说进化是每天都要面对的事实。读道金斯的书让我重新认识了信息在进化概念中的核心地位。我甚至在书的空白处写下“信息才是重点”。但当时我并没有彻底领会其意义。

后来故事的发展超出了我的想象。一旦理解了信息和进化的关联，就能够解释许多表面上没有明显关联的事物。现在我明白了，我们在日常生活中关注的大部分事物，用进化的信息可以得到最好的解释。结合了信息的进化策略的力量是如此强大和具有说服力，我开始认识到，复杂事物的形成机制应当作为我们整个教育体系所教授知识的基石。令人遗憾的是，直到2013年，在大多数大学的课表里还是很难找到讲述这两个概念之间的相互关联的课程。这本书的读者对象包括专家、学生和普通读者，同时我也希望它能推动创建新的创造性课程，探讨信息、计算

和进化对人类境况的意义。

我在这里给出的综合，任何受过教育的人通过广泛阅读也可以做到。不过，有引导的话会更方便。部分问题在于，要理解进化和信息的相互依存，人们需要用许多非计算机学家感到陌生的视角来认识世界。自学的另一个挑战是死胡同。我曾花费1年时间研究进化的热力学，最后发现这条路并没有什么助益。我最终认识到，进化过程从计算的角度理解才最有意义。一旦完成了观念上的跨越，对进化的热力学的理解就如同理解电脑为什么要插电才能工作一样自然。之前进化的热力学的神秘性消失了。另一个错误的方向是对“进化定律”的研究。我曾认为，进化过程应当表现出某种规律性，可以表示为简单的定律形式。150年来，许多极具才华的人都在寻找这样的定律。其中最大的成就是区分了达尔文理论的两大柱石，多样性产生和选择，并且阐明了多样性增加和自然选择的“定律”[4]。这并不能让人满意。从计算的角度认识进化有助于解释为什么进化定律如此难以捕捉[5]。

我发现有一条不遵循传统但极富成效的研究复杂性的途径，就是对两类复杂事物加以区分：一类是根据简单的化学和物理规律自发形成的简单无导向的复杂事物，例如岩石和星系；还有一类是具有目的性的复杂事物，例如橡树和喷气飞机，这类事物的形成依赖于复杂规则的存在；可以将这些规则视为“指令”。我认为，这些复杂的指令所包含的信息都源自进化计算。这个观点的意义深刻。

在草稿阶段，我曾认为我的论证有足够的说服力，那些因宗教或哲学原因排斥生命进化的人将不得不改变自己的思想，从逻辑上接受这个科学真理。现在我承认这不太可能，因为逻辑并不是障碍。达尔文理论的逻辑已经澄清了150年，然而还是有人拒绝接受它。对这些人来说，从更广义的角度思考进化，用统一的理论解释生命、免疫系统、人类思维、科学技术和现代经济，也许同样无法令其信服。人们排斥进化过程的解释效力和创造性的原因往往并不是基于科学逻辑，而是因为害怕。一些人害怕深刻的认识会有损他们的信仰，只会带来绝望而不是希望，对于这些人我想说的是，具有无限创造性的宇宙也许并不会那么让人沮丧。

还有其他作者也提出了生命的进化利用了简单的信息处理策略的思想。丹尼尔·丹尼特将这个策略称为“进化算法”[6]。道金斯强调“复制因

子”——能够复制自身的实体——的关键作用。而我强调的则是组织成指令形式的信息的重要性，以及包含了复制、变化和选择的进化过程（基于道金斯复制因子的丹尼特算法）是创造这种信息的唯一可行途径。

我给出的视角为理解世界提供了一条路径，这条路径最近才由物理、哲学等各领域的思想家编织成一个整体，还没有被广泛认可为一个正式的领域。这个综合依赖于计算机科学的深刻原理，但理解它不需要精通计算机理论。我没有列出数学公式，因为所涉及的数学十分抽象，会让大部分读者望而却步，而且这些数学对于理解核心思想并不是必需的。

在发展这些思想时，我惊讶地发现目的和创造性这样的人文理念也可以从计算的角度来解释和理解。我们将看到，目的可以通过重复的选择自然涌现出来，创造性则是许多进化系统的自然产物。我认为，本质上，所有创造性都来自源源不绝的“随机性”，更准确地说是随机选择。进化过程有效地利用了这个资源，一旦适当利用，一切皆有可能。如果没有见过章鱼或电脑，谁能想象它们的存在呢？

被认为最有效的科学认识世界的方法一次又一次发生根本性改变，这被称为范式转变 [7] 。由于科学和社会的联系变得如此紧密，这些转变会深刻影响人类对自身及其境况的认识。许多人包括我在内都认为科学和社会正处于这样的转变之中。过去三四百年间发现的许多科学原理都表述为描述宇宙特定方面规律行为的数学公式。牛顿的万有引力定律就是很好的例子。他给所有物体都赋予了质量属性，然后提出假说，认为所有质量之间都有被称为引力的相互作用力。万有引力定律提出，这种力的大小正比于两者质量的乘积除以两者距离的平方。这个关系简单而具有普适性。牛顿之后（他于1727年去世），科学家又发现了大量刻画物理世界规律性的简单关系。

看到数学关系的发现带来的巨大成就，人们很难不认为，只要发现了适当的方程，一切都可以描述。基于方程的世界观的一个后果是“决定论”。简单的方程严格可靠：填入同样的数字总是会得出相同的结果。但生命并不是这样，更不要说人类社会。简单的关系不能解释一切。许多现象都太过复杂，无法用简单的方程描述。没有方程能替代基于计算机的字处理软件，如果生命的历史可以重来，没有理由认为结果

会一样。恰恰相反，进化论认为，如果生命的起源和演化可以一次又一次重演，每一次的细节都会不同 [8] 。与此类似，与我们在物理学中期望看到的可预测性不同，人类的行为也具有不可预测性。这表明我们需要有新的认识和探索世界的方法。

越来越多各领域的学者认为这个新的方法就是信息。我们生活在信息时代。数字化无处不在：手机、电影、钟表、股市，你还可以举出很多。计算机让越来越多的活动变得更加迅速和可靠。设计汽车需要它们，没有它们飞机就无法起飞，智能手机和家用电器都使用微处理器。变化的不仅是这些物品，也包括科学认识的基础。量子力学被用信息重新解读。1990年，20世纪最著名的理论物理学家之一，普林斯顿大学的约翰·惠勒，提出"万物源于比特"，以强调信息在我们的物理宇宙中的基础性作用。他希望这句话能激发学术界重新思考怎样从最根本的层面认识宇宙。MIT的机械工程教授赛斯·劳埃德认为这条思路最有可能解决目前物理学的最大挑战 [9] 。生物学也越来越像信息科学。50多年前发现DNA在其化学结构中存储了生物的描述信息。这种信息是怎样被用于构造和维护生命体是现在许多生物学研究的重点。心理学家也在逐渐接受人脑的首要功能是处理信息的观点 [10] 。社会组织及其互动也越来越多地用信息和进化来解释 [11] 。

用一个词描述信息处理，就是计算。任何信息操作都可以被认为是计算，计算机科学就是研究信息操作的局限和潜力的科学领域。计算机科学在20世纪40年代成为一个正式的研究领域，当时在战争中需要建造计算机破译密码和计算弹道。今天这个领域在理论和实践上都发生了巨变。

从物理学到社会科学，几乎所有领域都从信息和信息操作的角度被重新审视。因为基于信息的观点经常能带来更深刻的认识，并解决一些难题。这本书中介绍了一些例子。我本人是生物学家，因此我举的许多例子都来自生命科学，但是我相信这本书同样也属于计算机科学和哲学。

书中每一节的标题都是一个问题。这反映了科学的特性，从根本上说，科学就是问问题和回答问题。这能让我的写作更专注，希望对读者也是如此。

最后是免责条款。在向非专家读者介绍学术尤其是数学概念时，必须简化，直指概念的本质，而不能纠缠于技术细节。如果不具备背景知识，通常很难正确完整地理解专业概念。因此，肯定有些专家会质疑我的阐释。这时他们应当问一问，这里给出的简化是否具有误导性，以至于还不如不向读者解释？或者只不过是缺少细枝末节？

第1章　问题

地球有多特别？

假设你是一个研究地球的外星科学家。你通过仪器发现，大气层完全背离化学平衡，氧气偏高，二氧化碳偏低。与附近的行星比起来，金星的二氧化碳浓度要高20000倍，并且没有自由氧。地球还发射各种波长的电磁信号，有些地方晚上还发光，而地表岩石的温度又没有到能发光的程度。拉近一点观察，发现有细长的带状物，长达数千千米，还有摩天大楼，以及植物组成的生态系统。再近一点，发现数以亿计的功能生命体。许多植物能产生极为复杂的化学反应，利用太阳光的能量合成各种化合物，释放氧气。动物则吃植物或动物，呼吸氧气，并进行同样复杂的连锁化学反应，将植物化学结构转化为动物化学结构，并暂时储存活动所需的化学自由能。还不仅如此。其中一种十分成功的动物还创造了结构复杂的非生命体，能极大地拓展其身体和心智能力。这个物种为了自身的利益改变了自然环境。显眼的道路和摩天大楼就是改变环境的两个例子。这个物种还发展了复杂的通信方式，并且创建了复杂的社会组织。一个星际旅行的外来观察者会马上意识到这繁盛的复杂性表明这个星球已经出现了进化系统。为什么她会得出这样的结论呢？这就是这本书要回答的问题。

人类已经发射了航天器探测金星、火星、木星和土星的卫星土卫六，并且在地球上和太空中架设仪器研究行星和它们的卫星。除了水星，所有行星都有大气层；除了月球，所有大一点的卫星也都有大气层。只要不是表层太软无法支撑结构，我们观察到的行星都有山脉和火山。有一些卫星还覆盖有厚厚的冰层，木星的卫星木卫一上还有硫火山。科学家发现了各种陌生或熟悉的现象，但是都能通过已知的化学和物理定律在一定条件下的长期作用得到解释。可解释事件的链条可以一直回溯到时间的开端，但如果某个事件所需的条件产生的概率极低，就无法用这种方式解释。不可能的事件一直在发生。目前科学还不能解释一切，但没有理由认为，奇迹或概率极低的事件能在宇宙历史中的某时某处发生。

地球不同于我们研究过的其他行星。虽然我们所知的世界里没有什么违反科学定律，但化学和物理课本却不能完全解释iPod、花草和贝多芬交响曲。化学和物理定律允许它们的存在，但却不能预测它们的出现，这些事物通过随机组合出现的概率低得难以想象，然而它们却随处可见。

显然生命的存在改变了概率的计算，但这是怎么做到的呢？生命似乎能创造无穷无尽的特征，这些特征无法用化学和物理定律进行预测，是什么赋予了生命这个能力？生命的存在又是如何预示摩天大楼和计算机的出现？光合作用和芝加哥西尔斯大厦有什么共同的原理是火星上没有的？又是什么让我们将一些事物，尤其是人类的天才创造，与非生物自然的产物区分开来呢？

为了富有成效地探讨这些问题，我们需要特别的视角。首先我们需要将事物分成更小的部分。这并不新鲜。希腊原子学派哲学家在公元前5世纪就认为所有事物都是由微小粒子（虚空中的原子）组成，而各种物体则是由原子随机组合而成。虽然时间久远，这个思想的一部分却仍在科学和哲学中回响。

现代观点认为物体是由原子和分子组成。组分的随机组合肯定会创造出许多不同的东西。如果有合适的组分，通过随机组合，经过足够长的时间，应当能产生出任何能想到的事物。从这个意义上说，随机蕴含万物。但很久以前柏拉图和亚里士多德就认识到，困难在于不受约束的随机很少能产生有趣的东西。如果组分的数量很多，那么任何特定形态的形成都会极不可能。将事物毫无意义地组合到一起的方式是如此之多，以至于要想见到特定的结构自发出现，等待的时间可能比宇宙存在的时间还要长亿万倍。

我们倾向关注的事物大部分都具有复杂的历史，并表现出具有高度针对性的复杂特征。现实世界中的许多结构让人明显感觉到具有目的性。它们似乎是为了某种目的设计出来的，或者为了适应某物，或者至少是在某种背景下“具有意义”。这引出了另一个问题：目的性有科学的解释吗？答案是：进化过程会自然产生出我们认为具有目的性的特性。

在关于进化的科学争论中，有一个重要的维度经常被生物进化的反对者忽视，就是现实世界的复杂性并不局限于生物。我们的周围到处都

有非生命的复杂事物，从交响曲到太空飞船，它们都无法用物理学和简单概率进行解释。生物和人类公司制造的大多数产品都无法仅仅用牛顿力学和量子力学得到完整的解释。还需要其他思想：具有高度针对性的结构的形成需要有被正确组织和使用的信息进行引导，否则就不可能出现。正是对这个原理的有意识利用使得支撑现代文明的所有技术进步成为可能。

与之相应的一个原理是：生命和技术性的复杂事物显然的不可能性反映了形成它们所需的附加信息的量。所有生命和人类创造的非生命产品都体现了这一点。形成某物所需的信息越多，其表现出的不可能性就越大。举个例子，比较石斧和超级计算机，或者病毒和人类。越复杂的因而也越不可能的事物，制造它们所需的信息也越多。这种描述信息的形式表现为蓝图、配方、基因（DNA或RNA），或者统称为指令。指令可以有各种形式，但一个简单的定义可能是这样："指令是特定的编码信息，具有以下属性，当用来依序改变另一个系统时，可以得到预定的最终结果，并且对于创造这些指令的某人或某物来说，能带来某种好处或所希望的东西。"指令含有目的性信息。它们被创造出来是因为其产物能给创造者带来利益。指令通常指定或预先指定另一个系统需要执行的步骤和行为。这个系统可以是人类，也可以不是。步骤的完成会导致某个特定目标的达成。因此所有指令都隐含有目标。

这些指令一旦被执行，就会产生某种事物或行为，如果没有这些指令，这些事物或行为随机产生的概率极低，基本没有可能性；而有了正确的指令和执行指令的系统，产生出的结构就可以被认为是有可能的。显然要理解生命和人类天才的创造，我们必须解释指令的来源。

前面给出的指令定义很抽象，很难让大多数读者接受。举个例子会容易理解些。后面还有更多例子。

指令的基本思想大家都很熟悉，一个例子就是食谱。蛋奶酥这种食物只有人类会做。要做蛋奶酥，需要遵循一系列精准而简单的步骤。只要遵循了这些步骤，就会得到特定的结果。下面是网上找的一个蛋奶酥的做法[1]：

1.将6个鸡蛋分为蛋黄和蛋清。

2.溶化1/4杯奶油，混合1/4杯面粉，1勺盐，1/4勺芥末和少许辣椒粉。加热搅拌2分钟。

3.缓慢加入1杯半牛奶，持续搅拌直至匀滑。再加热2分钟，持续搅拌，然后停止加热，加入1杯碎黄奶酪搅匀。

4.将蛋黄打散，加几勺热奶酪酱，加热蛋黄，然后将热蛋黄加入奶酪酱充分搅匀。

5.用电动搅拌器将蛋白搅至沫状，将1/3蛋白加入奶酪酱搅匀。充分搅匀后，将剩下的2/3加入奶酪酱。

6.将奶酪酱倒入2升平底烤盘，放入预热的烤箱200摄氏度烘烤12-15分钟。

照这个食谱小心烹制就能做出美味蓬松的佳肴，远胜于你将这些配料直接混在一起在锅里炒出的味道。奶蛋酥要比炒蛋复杂得多（也难做得多），因而也更特别。特别在哪里？我认为其特别之处可以用食谱的复杂度衡量。从某种意义上说，与炒蛋相比，蛋奶酥所需的步骤更多，因此也更复杂。在这个例子中，复杂度增加不多，我们常常会遇到还要复杂得多的事物。即便是这样，就算是最有创意的厨师，如果没有看过食谱或没看别人做过，也很难做出蛋奶酥来。这个描述指令是1维和线性的，其他可能还有并行（同步）的步骤，还有一些是2维的，比如蓝图。无论采取哪种形式，它们的编码都是由某人或某物执行的特定的具有目的性的信息。

但并不是所有事情都需要指令。我们可以将周围无穷无尽的事物、系统和行为分为两类。一类包含那些形成需要指令（或配方、蓝图）的事物；也就是说，除了化学和物理定律以外，它们还需要附加信息。我们姑且称这些需要指令的事物为Ⅱ型事物。这其中包括所有生物，以及大部分人类智慧的产物，包括蛋奶酥和宇宙飞船。还有一类则包含所有无需指令的事物，我们称之为Ⅰ型事物。这一类包括大部分非生命或非人类活动的产物，例如岩石、海洋、气象和太阳系。两种类型一起涵盖了一切事物。

将事物分为两类可以将注意力集中在一些（但并非全部）形成过程中需要附加信息的事物，同时也回答了这一节的问题。地球之所以独特

正是因为Ⅱ型事物。所有已知的Ⅱ型事物要么是在地球发现的，要么就是在地球制造的，Ⅱ型结构将地球与其他已知行星区分开来。当然，也可能还有其他未知行星也有Ⅱ型结构。地球上也有许多Ⅰ型结构，但这一点并不能将地球与太阳系或系外的其他行星区分开来。生成Ⅱ型事物所需的指令编码了目的性信息。指令总是具有意向性。它们有固有的目的性。因此，我们面临着3个任务，认识“信息”是什么，了解指令的来源，以及研究指令固有的目的性来自哪里。最终进化计算将完美地解释指令的来源和目的性。

总结一下，从化学和物理学的角度看，世界上的许多事物是如此复杂和不可能，以至于很容易认为，如果没有超自然的智慧力量干预，它们的存在是不可能的。指令提供了另一种无需求助奇迹的解释。

什么是信息？

我们经常听说现在是信息时代，你可能和我一样，也对信息的用途和意义感到困惑。手机、电视机和计算机都是“数字”的。它们的功能依赖于编码为0/1序列的信息。这些序列编码了信息。将0和1组成的序列输入特定的设备可以产生出我们想要的东西。随便翻开一本词典，都能看到将信息与获取知识关联起来的信息定义。对于信息有一种认识，即它必须对某人或某物是有用和有意义的。领导计算大原理[2]计划的彼得·丹宁和克雷格·马爹利这样定义：“信息是用符号模式传递的意义。”这个定义相当直接，也体现了信息的另一个重要特性：那就是可以被编码，能通过各种媒介存储和传递的消息。在计算机科学以外的领域，学者们通常将信息等同于获取知识。

信息的数学定义也很有用，但是无法体现上面这种意义。利用数学定义，我们可以计算呈现在特定场合（例如特定的符号序列）中的这种东西（信息）的量，但这无法告诉我们它有多少意义，或者一个符号序列有什么用。不过如果我们知道所研究的信息是如何产生，就可以改变这个局面。对于进化系统，或者推而广之，对于能学习的系统，获取的信息必定刻画了所适应的或学习的东西的某个方面。因此如果认定信息存在本身就是为了表达某种意义，就可以在一定程度上量化意义[3]。

量化定义是信息革命的基础，因此有必要了解一下数学家和工程师

眼里的信息是什么。我们先了解一点历史。19世纪末，科学家们开始研究由原子和分子组成的粒子世界，并将这个新的视角与已经建立的热力学联系起来。古典热力学的研究对象是“系统”：边界内的物质组合。系统可以是复杂的机器，或是烧杯中的化学反应，甚至可以是房间里的空气——各种有边界的东西。所有系统都用能量、熵和物质刻画。热力学研究的是系统变化时能量和熵如何变化。热力学第一定律认为能量可以转换形式，但无法创生或湮灭。一些转换（但不是全部）可以做功。热力学中做功泛指与能量转换有关的活动。比如抬箱子、开车或是创造新的化学结构都与做功有关。在做功过程中，能量守恒，但是会转换成做功能力较差的形式。熵可以刻画一个系统做功的能力。系统的熵越大，做功的能力越差。热力学第二定律认为封闭系统的熵要么增加，要么不变，永远不会减少。这条定律很容易引起误解，因为我们平时遇到的系统很少是封闭的。当封闭系统的熵达到最大，我们就说这个系统达到了平衡态，不能继续做功。一般来说，系统会变化直至达到平衡态，然后就不再变化，除非某种外界影响作用于它（对系统做功或者增减某种东西）。

行进的汽车就是系统朝平衡态变化的例子。如果你改装一辆汽车，用气罐为发动机提供空气，用绝热的袋子收集发动机产生的所有尾气和热量，你就能有一个大致封闭的系统。这辆改装的汽车可以开动直到汽油（或空气）耗尽，然后就不能再移动，这就是处于平衡态。开车过程中系统的总能量不会改变，但汽油和（气罐中的）氧气会逐渐转化为绝热袋中的二氧化碳、水和热量。一旦汽油（或氧气）被转化完，汽车就无法再开动，除非你重新加油或补充压缩空气。将二氧化碳、水和热量放回油箱和气罐没有用；汽车需要汽油和（空气中的）氧分子提供的低熵形式的能量。一旦转化成二氧化碳、水和热量，系统的总能量不会变，但系统中能量的表现形式会退化成无法继续做功推动汽车的状态。二氧化碳和水有可能还原成汽油和氧气，但这需要消耗外部提供的低熵能量。

熵的概念可以从分子层面理解，德国物理学家玻尔兹曼和美国化学家吉布斯提供了思想和数学基础。根据这种思想，物理系统是由大量原子或分子组成，系统的状态可以用所有粒子的瞬时位置和速度来刻画。系统的所有分子在某个时刻的瞬时状态被称为微观态。由于分子不停地运动，系统也就不停地从一个微观态转换成另一个微观态。（我们看到或测量的）可观测状态实际上是测量期间许多微观态的时间平均。根据

这个思想，熵实际上是系统状态所对应的微观态的数量；对应的微观态越多，熵就越高。也就是说，熵是对不确定程度的度量。熵越高，系统具体处于哪个微观态就越不确定。系统无法自发到达一些具有结构（比无结构系统的熵低）的微观态。

20世纪40年代和50年代，香农将物理学中熵的概念应用于通信。要解决的问题是如何设计一个系统可以在指定时间内传送所需数量的信息。为了解决这个问题，必须量化系统传送信息的能力。香农的基本思想由两部分组成：一个系统如果可以有多个状态，就具有携带信息的潜力；而且，可能的状态越多，一个具体的状态所含的信息就越多。香农的度量方法是基于概率。因为在给定条件下，所有可能状态的概率加起来必须等于1（必定有一个发生），可能的状态越多，平均下来任何一个特定状态发生的概率就越低。

举个例子，“the”这个词由3个字母组成，用t、h和e组成3个字母长的词有27种方式（例如thh、the和eee）。首字母为t的概率为1/3（必须是t、h和e中的一个）。同样，第2个字母为h的概率也是1/3。根据概率原理，如果事件是由子事件组成，每个子事件有独立的概率，则组合事件的概率是子事件概率的乘积。因此从一个有相同数量字母t、h和e的袋子中顺序取出3个字母刚好组成词the的概率是1/3×1/3×1/3，即1/27。再来看evolution这个词。它有9个字母长，使用了8种符号。用8种符号字母组成9个字母长的序列有8×8×8×8×8×8×8×8×8种方式（科学记数法为8^9 ）。结果为134217728。因此从装有许多字母e、v、o、l、u、t、i和n的盒子中取出9个字母刚好组成词evolution的概率为1/134217728。如果你想自己尝试，得多准备点时间！随着可能序列的长度和可能字母的数量增加，可能的序列数量增加得非常迅速。从全部26个英文字母中随机取字母组成evolution的概率是1/26×1/26×1/26×1/26×1/26×1/26×1/26×1/26×1/26，大约为5万亿分之1（12个零）。这个句子由118个字母、空格、数字和标点组成。26个字母、2个标点、2个数和1个空格（总共31个符号）组成118个符号的序列大约有1后面跟176个零这么多种可能。这是一个非常大的数，比可见宇宙中的原子数量还多，这还只是一个句子！想想整本书的可能性。

通信需要从可能的序列或状态中识别出1种。从通信工程师的角度看，可能的状态数量越多，确定下来的具体状态携带的信息就越多。文字对交流很有用，因为每个文本都十分独特——有许多种可能。一个很

简短的文本也会携带很多信息。但是通信并不关心文本的意义，只关心它有多特别。

可能状态的数量、概率、文本的长度以及文本所携带的信息密切相关。长文本携带的信息比短文本多，长文本的概率也低得多。香农设定信息的度量是可加的；换句话说，如果两个文本的信息量分别为X和Y，则组合在一起的信息量应为X+Y。长度具有这个特性，但概率没有。结果发现这个问题本质上与玻尔兹曼面临的数学问题是一样的。区别在于信息系统是由符号序列组成，而物理系统则是由空间中移动的分子组成。让人难以置信的是，这两个看上去截然不同的情形却有着相似的数学框架。香农天才地认识到了这一点。

香农提出的对文本信息量的度量首先要给出所有可能状态（系统允许的每一种字符排列）的概率，并与概率本身的对数相乘，对每种可能的符号序列都计算这一项，然后将所有项相加。注释中给出了数学表达式 [4] 。如果对数以2为底，则所有项之和就是系统的香农信息量，单位为比特。

这个度量有两种解读，取决于你的视角。从通信工程师的角度看，香农度量刻画的是系统携带信息的能力。这种观点关注的是不确定程度——系统的这个量越大，你对于文本所知道的就越少。因此香农将他的度量称为熵，与热力学中的术语一致。

从另一个角度，即你感兴趣的具体对象的角度看，香农度量可以被视为这个对象所携带的信息量（不一定是通信）。这也是这本书大部分时候采取的视角。容易让人混淆之处在于，如果你得到了具体的对象（字符串），不确定性就消失了，（从通信的角度）香农熵为0！然而，如果将具体对象放在所有未实现的可能状态的背景下考虑，则可以认为具体对象所携带的信息量等同于对象所属系统的熵 [5] 。有时候也直接将对象或事件概率对数的负值定义为香农“自信息”。

（两种意义上的）香农定义的一个特点是，如果只使用2种符号（0和1），并且所有可能符号序列的概率都相等，则信息的比特量就等于序列（或对象）的长度。也就是说1000个0和1组成的字符串（例如10110100010……）携带的信息量是1000比特，除非有某种规则使得序列中1和0的概率不相等。这个例子揭示了香农信息度量与序列长度的密

切关联。

如果状态的概率（注释4中给出的那些）不一样，则携带的信息量低于能够携带的最大信息量。不过概率取决于背景。它们取决于现行的规则或已知的是什么。一般可以认为不具有背景知识，直接假设各种可能的概率相等。这时一个对象的“信息量”就等于（以这个对象为具体状态的）系统的最大熵，也等于（表示为二进制的）序列的长度[6]。为了简单起见，我通常用序列的长度作为序列可能携带的信息量。当信息表示为二进制时，长度单位为比特。这也是描述DVD容量或计算机中文件大小的度量。

香农的公式是现代通信的基础，但还有一些值得注意的地方。例如，当用于刻画对象时，它没有考虑冗余（概率原因的冗余除外）。因此，重复消息“I went to the marketI went to the market”包含的香农信息是单句的2倍。从常识来说这很愚蠢。如果你读了什么东西然后记住了，再读一次不会让你的知识翻倍。还有一个容易混淆的地方是香农信息度量与意义无关。“wtm naIt oeekh t tre”包含的香农信息与“I went to the market”一样，因为有一样的长度，用的也是一样的字母。香农度量严格基于系统可能编码的不同信息的数量和概率，与其他的无关。有一个调和日常意义的“信息”与香农信息的简单办法是将香农信息视为文本可能包含的最大信息量。实际的意义或用途可能比这个量少，甚至可能为0。

根据物理学，当分子可以在空间自由运动历经所有可能的微观态时，分子系统具有最大熵。当存在结构时，一些分子无法四处自由运动，它们被限定在特定的位置。因此在相同的温度条件下，如果原子数量和种类相同，具有物理结构的系统比不具有结构的系统有更低的熵（更低的自由度）。实际熵与抽象的最大熵之间的差值有时候被称为负熵。

符号序列中也可以存在结构。这时的结构指的是符号的非随机分布。符号不会在序列中移动，非随机指的是概率不相等。如果可以确定具体的概率，用它们计算香农熵，然后与（概率相等的）最大熵相减，就能得到与物理学中的负熵对应的信息量。这个量度量了存在的结构[7]。

任何由多个部分组成的系统都可以用熵和信息刻画。重要的是，还有其他与信息类似的刻画系统的方法。例如，当考虑有明显内部结构的系统时，很显然一些比另一些要更复杂。结构的这方面性质接近复杂性，却没有被香农信息、熵或负熵体现出来。后面我们还会对此进一步讨论。

什么是进化系统？

所有进化系统都是基于存储的信息，并具有完整的信息处理策略，能够对信息体进行修改、添加和删除操作。所有进化系统的核心都利用概率计算，可以从随机发生的事件中提取目的性信息。一旦这种计算有可能生成指令，再用指令生成有用的事物，就有可能形成正反馈循环。指令的变化如果能改善结构或行为，就有可能以此为基础在将来进一步改进指令及其产物。通过这种高效的循环，指令不断改进。当指令是用于解答某个问题时，如果最初对答案是怎样的没有任何想法，这种策略将是寻找答案的最有效的方法。

当从信息的角度来认识相关的对象和活动时，会发现随机性才是新事物的终极来源。对计算机科学不熟悉的人可能会觉得难以置信，但其中的逻辑是合理的。确定性的规则精确决定结果；因此只要输入不是随机的，结果就一定可以预料。确定性的计算机之所以如此强大就是因为只要给它们同样的输入，总是会得到同样的结果。这是创造性的对立面。创造性意味着意料之外的结果，但在确定性过程中这样的结果只有当输入不可预测时才有可能。不可预测性是随机性的精髓。所有创造性活动的设计都隐含了随机输入，这本书中讨论的进化过程几乎都可以视为从随机来源进行提取的概率计算。

所有进化过程的共同特征包括以下5个要素[8]：

1.个体。一般都有许多某种类型的单元体。它们有各种名字，比如：生物、自主体、基因、概念和公司。

2.可遗传的特征。个体的描述信息，以某种形式编码为个体本身的一部分（生物的这种信息编码为DNA）。

3.个体可以繁殖或复制。通过这个机制，个体从父辈或之前的个体

拷贝编码信息。

4.变化机制。信息在复制、繁殖或维护过程中必须有机会产生适度的改变。在许多系统中变化机制就是复制过程中产生的错误。

5.基于特征的选择。繁殖（或复制）的成功必须部分取决于各个体编码信息所描述的特征。

只要系统同时具备了这5个要素，个体组成的群体中的编码信息以及相应的个体特征就会随时间改变；遗传的个体特性必然会越来越适应决定繁殖（复制）成功率的标准。

具有这种逻辑结构的系统会累积适应选择标准的编码信息。它们利用随机变化做到这一点，有时候也利用非随机变化，以免偶然性过大，不利于产生有用或有趣的东西。这个信息累积和改进的过程就是概率计算。它有效的原因很简单。很小的变化通常是有可能的。最小的可能变化是是/否、增/减、要么/要么这种类型的。这样的选择一般有一半对一半的概率发生，也有一半对一半的概率一个比另一个好。而无目的的大变化基本不可能发生。通过小的并非很不可能的变化，并累计好的变化，就有可能达到本来很难达到的目标。著名的进化论作家道金斯有一本书的书名就体现了这个思想：《攀登不可能的山峰》[9]。

可能会发生错误的指令复制机制一旦与在多个结果中选择的机制相结合，就会产生戏剧性的变化。这种选择复制过程一旦成型，即使没有引导，也会逐渐装配出本来出现概率极低的复杂目的性指令。复杂的指令反过来又使得本来极为不可能的复杂对象或行为的形成成为可能。后面我们会看到化学和物理定律作用于自然产生的物质也可以自发形成结构，有一些还相当复杂。这类对象的形成不需要指令。无需指令的结构的形成局限于有较大概率自发形成的事物。而指令则为创造有用的、神奇的、聪明的、美丽的概率极低的事物提供了无限可能。在第4章我们会看到一些例子。可以毫不夸张地说，不违反化学和物理定律的任何事物都可以利用适当的指令得到——相当惊人的论断。当然，挑战在于得到正确的指令。没有指令，可能性就会大为受限。

由于指令的使用是如此强大而普遍，最好是将使用指令产生的复杂性（Ⅱ型事物）与不需要指令的复杂性（Ⅰ型事物）区分开来，并且将

Ⅱ型事物归于更高形式的复杂性。我在书中反复强调，所有最复杂的指令都是由进化计算创造的（我们不知道还有哪种计算策略能做到这一点）。对这个断言的一个检验是只要发现了不一般的复杂性，就应当能识别出明确的编码指令以及创造指令的计算的物理基础。对于生物，信息是用DNA分子编码。细胞和生物通过生存和繁殖执行这种计算。实际上，生物依赖于两种非常不同的计算。一是创造和维持生命本身。这些过程部分依赖于DNA编码的信息[10]。不是所有人都将这些化学和发育过程视为计算。另一种很不一样的计算是基于DNA的信息一代又一代慢慢改变。这才是这本书所关注的进化计算。名为群体遗传学的生物学分支描述的就是改变生命的进化计算是如何一代又一代进行的。

确定性的电子计算机也能够执行进化计算，但与生命执行的不一样。对于计算机，进化的信息是软件代码。大脑也可以被视为是计算机，信息则编码为神经元的通信模式。在第10章，我们会探讨学习和思维活动是进化计算的可能性。

从信息累积和提取的角度看待进化可以带来更深刻的认识，同时也让一些原本很困难的问题变得容易解决。例如，进化进步的观念在生物学中存在争议，也没有人否认生命的最终起源是一个尚未解决的科学难题。计算的视角有助于澄清这些问题。对于进化进步，如果进步等同于复杂性的增加，则根据信息理论可以确认，特定情形下复杂性的不断增加是可以预期的结果。后面我们会探讨复杂性的另一个方面——深度，并证明为什么它在特定条件下会增加。书的最后还会探讨为什么所有进化系统的起源包括生命在内都要受信息理论的限制。总的来说，由于进化计算具有产生越来越大的复杂性的巨大潜力，有理由怀疑只要出现了显著的复杂性，就存在进化引擎。

什么是科学的认识？

我写这本书的目的是希望能让读者更深刻地认识世界，包括我们自己。我尽量让自己的解释可信，让读者满意。结论可以通过观察和理论的逻辑推导得到，但只有通过实验检验才会被接受。毫不奇怪，这正是科学研究的方式。读者需要自己评估论证和检验结论是否可信。

从根本上说，科学之所以可能是因为宇宙有可预测的模式。科学的

中心任务就是发现这些模式。不是所有事物都有规律，许多都没有，但规律肯定在某处存在。牛顿的万有引力定律（具有质量的物体有相互吸引力，力的大小反比于距离的平方）就是这样的规律 [11] 。在牛顿之前，虽然所有人都知道物体下落，但没有人认识到其本质。

但科学的认识并不仅仅是观察到的一系列规律。要认识某种结构或活动，我们不仅要知道相关的定律，还要能预测它们在特定情形下如何随时间变化。例如，工程师要能利用引力定律、能量守恒和动量的概念预测炮弹的轨迹或坐过山车时感受的压力。一旦预测得到验证，我们就能对所采用的原理有信心。科学之所以如此受推崇正是因为每天都有许多基于科学的预测得到证实。

装置能阐明原理和验证认识。如果借助科学认识创造出了能产生可预测结构和行为的装置，认识就能被证实和丰富。以声波理论为例，理论认为声音由波组成，通过媒介（通常是空气）分子的密度变化传播。扬声器和麦克风正是基于这个原理。扬声器的弹性表面可以被设备（通常是电磁铁）前后驱动。表面的快速运动会压缩和舒张附近的空气，从而形成运动的压力波。但我们怎么确定这种波的确存在呢？麦克风可以用来检测压力波，压力波会改变麦克风弹性面的形状，弹性面的运动转换成电信号，通过放大就能驱动另一个扬声器，或是在示波器上显示，也可以记录下来等将来使用。如果声波理论不正确，扬声器和麦克风的设计和实现就是不可能的。人类制造的这些装置阐明和证实了科学理论。

科学提供了两个相容但又大致独立的对事物的解释——物理和生物进化。物理主要是解释非生命现象，但也能用来认识生命功能。生物进化解释生命的起源以及它们如何适应环境。

传统进化论对于理解星系或原子不起作用。不过近50年来进化的解释范围逐渐拓展，开始涵盖生命以外的事物，包括人造物和特定类型的计算机程序 [12] 。一些物理学家甚至提出宇宙的演化可能也与生物进化类似。

300年前，牛顿用几条简单的原理和微积分这门新的数学工具，用一种全新的方式解释了太阳系的复杂性。在此之后，科学的巨大进步带来了新原理的发现，并据此重新解释了我们的物理世界。今天的生物化

学家、分子生物学家和生理学家就是利用化学和物理定律解释生命。

150年前，达尔文用全新的方式解释了生命的复杂性。许多科普读物都指出了进化是理论，而物理则是基于定律。在科学中，理论是对现象的解释。理论可以是猜测性的，也可以是严格确证的，取决于验证的强度。定律一词在科学中的定义并不完善，但定律也可以是猜测性的或是严格确证的。定律主要是描述关系或原理。定律通常可以表述为一个简短的命题或简单的方程。牛顿万有引力定律就是两个质量块之间的关系的代数表示。热力学第一定律（能量既无法创生也无法湮灭）则表示为原理。

进化既不是关系也不是简单的原理。进化是一个过程，基于之前的成果创造更先进的成果。因此，进化很少表示为定律，虽然它是强大的并被广泛认可的科学解释。同物理过程类似，进化论也是基于测量、数学定理和物理定律。进化并不是唯一解释过程产物的理论。星体的形成、大陆漂移和工程师解释的内燃机工作原理也是过程理论。成功的解释性理论要基于物理定律和细致的测量，但它们的内涵通常无法用简单的方程表示。

与一些反进化论者的论调相反，科学并不仅仅限于只能在实验室里直接验证的问题。经常有一些现象，虽然人类还无法复现，但还是被认为已得到充分认识。例如山脉、星体和恐龙蛋。对物理现象的解释只要遵循了已被认可的理论、原理和定律，并且可以构想、模拟或示范出合理的机制，就会被认可。很久以前起源的生物和生物特征的进化论解释就是这样被认可的——还有许多太大或者形成过程太慢以至于无法在实验室复现的事物的科学解释也是一样。

有批评意见认为进化论没有为重要的进化产物如何通过逐步的小变化产生提供明确的证明，例如能调焦的眼睛、细菌鞭毛或细胞代谢的复杂性 [13] 。但是许多论证后来被发现是基于误解或是对科学文献的理解不完整。近20年来，在理解调节基因如何塑造动物的器官和附器方面的进展迅速。新的发现逐渐澄清了进化过程是如何通过合理的小进步逐渐创造出全新的器官和生物。新出现的学科进化发育生物学就是从分子机制的角度研究如何通过随机突变和自然选择的累积变化创造出各种生物结构。这个过程的关键在于已有的结构同时也可以有新的用途。有一个早就知道的明显例子：鱼鳍逐渐演化成用于行走的腿，前腿演化成（蝙

蝠和鸟的）翅膀，颌骨演化成内耳骨，叶子演化成（沙漠植物）保护性的刺。最近，DNA序列分析发现，今天对复杂器官的形成很关键的高分子在远古时期曾对一些似乎无关联的结构的形态改良起作用[14]。

虽然还没有人能用非生命物质人工创造出活细胞，或用进化的方法创造出眼睛或耳朵之类的复杂结构，但现在几乎所有生物学家都承认进化论是对生物复杂性的正确的普适性解释。无数实验证实了选择具有改变从病毒到奶牛等各种生物种群的力量。没有进化论，就无从解释生物对环境的适应，以及各物种DNA之间的关联。成功的现代农业和养殖业养活了全世界，也验证了进化论的正确性。可以想见如果对动植物的人工选育持续百万年会带来怎样的成就。

科学的认识带来的不仅仅是对现象的解释，同时还有难以想象的技术进步。其中许多已经改变了人类的生活和我们居住的星球。如果你不相信，想一想400年前的人口还不到现在的1/20。当时的人口受制于疾病和饥饿，大部分人的预期寿命不到30岁。大多数人从事农业，夜晚靠月光和油灯照明。出行靠步行、马或帆船。远距离通信是少数人的特权。今昔对比，差别主要在于科学技术。当然社会组织和社会结构的进步也改善了许多人的生活，但大部分改变都是因为新的技术才有可能，没有人能否认过去400年科学技术的进步彻底改变了世界。

人类远在有文字记录以前就在创造技术，但大部分时间进步都很缓慢。是什么加快了创新的步伐？答案很简单：一旦认识了相关的自然规律和原理，就能利用这些原理发明新的装置。没有对万有引力和推进火箭的化学反应的认识，就不会有探索火星的宇宙飞船。自然界不存在像航天器这样能依照我们的想法行事的结构复杂的装置。科学解释存在的现象，技术则利用这种认识创造新的装置。因此，技术成就证实了背后的科学认识的正确性。如果麦克斯韦电磁理论不正确，设计的发电机就发不出电，墙上的开关也打不开灯。我们每次开灯都是在证实对电的科学认识。如果科学认识是错的，工程师就无法依靠它建造能工作的设备。这就是为什么占星术、炼丹术、巫术和宗教教条对于建桥和治病这类实际事务没用，以及为什么它们不具有科学的可信度。再强调一次，没有科学也可能有技术，但是会很有限；而一旦受科学引导的技术有效，它就证实了其背后的科学。

那么对于什么是科学的认识这个问题有没有简单的答案呢？有，科

学的认识是所有透彻研究和学习了某个问题的人们当前的共识。这样的共识总是可以通过设计新的实验来探索和检验，以确定这种认识的逻辑推论是否与新的结果相符。如果不相符，就是哪里出了错，要么是实验的逻辑需要被重新审视，要么是基本的认识需要被修正。反过来说，如果有人想挑战目前的科学认识，他们所需做的就是用逻辑分析或可重现的实验证明得到的结果与目前的认识不相容。然后学术界就会去重现这个实验或是确证其逻辑，如果新的实验可靠，原来的认识就会被修正。通过这种方法，科学认识就会不断改进——或者说进化？

计算机科学如何加入？

伽利略之后，数学成了刻画自然规律的通用语言。我们对物理现象的认识大部分都被写成数学语言，数学也被广泛用于预测将来的结果。生物学和社会科学似乎没有遵循这项惯例，这通常被归因于它们极端的复杂性。许多人都认为找到合适的数学支撑生物学或社会工程只是时间问题。我的观点是合适的数学已经找到了，就是计算机科学。

计算机科学及其相关的理论可以帮助认识传统数学对于理解复杂过程的作用。计算机输入的是程序和数据，程序操作数据中包含的信息，操作的结果称为输出。

物理世界可以用类似的方式理解。如果将物理定律视为宇宙的程序，物理结构就是输入输出。雪花是化学和物理定律作用于温度低于凝点的含有过量水蒸气的空气（输入）的结果（输出）。广义上我们可以将物理过程视为计算，自然规律作用于输入状态产生输出状态。

计算机理论所揭示的输入输出的关系对于理解各种复杂事物很重要。一个非常重要的原理就是，简单的规则作用于简单的输入可以产生极为复杂的事物。当输出作为下一轮计算的输入时尤其如此。有了合适的规则，重复过程运作得越久，输出就会变得越复杂。一个典型的例子是通过简单操作产生复杂图像的计算，反复用每一轮的输出作为下一轮的输入。图1.1展现了用这样的过程得到的图 [15] 。物理宇宙中的许多事物也可以通过反复应用简单规则得到。雪花、星系和火山就是简单物理定律反复作用的产物。

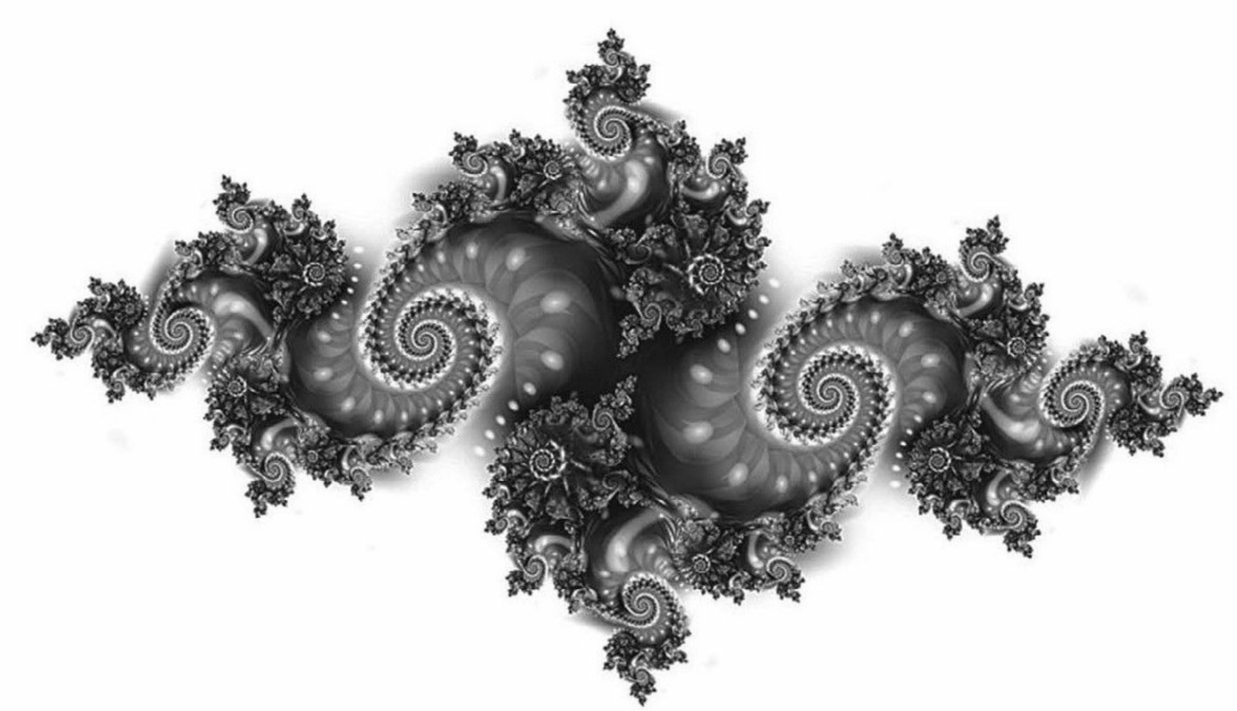


图1.1　计算机反复应用简单规则生成的图形

计算机理论中还有一条不那么明显但同等重要的原理是，一些复杂事物无法通过应用简单规则得到，无论过程执行多长时间。这类事物只能通过应用复杂规则即长程序得到。由于化学和物理定律，以及传统数学中的大部分，都是处理相对简单的关系，通过计算机理论我们可以知道，应当有一些结构的形成无法表述为简单的数学关系。现代计算机技术提供了大量这样的例子。想一想用来创造这份文档的计算机程序。微软的WORD提供了各种服务，包括保存、格式、查找、语法检查和拼写检查。告诉我的计算机该如何做这些事的基本程序文件占据了8.4　M硬盘空间。因此，将我敲的字转化为格式化文档的“规则”比这本书（大约1　M文本）还长。越简单的字处理软件程序越短，但没法缩减到只有几条规则。没有哪个数学分支能为文档的格式化计算提供工具。同厨师烹制蛋奶酥一样，格式化最简单的方法是遵循一长串规则。之前给出的食谱就是现成的例子，不可能为蛋奶酥写出简单的做法。

如果输入和输出之间的关系可以用数学表示，输出就往往是可计算的。数学的威力在于提供捷径。如果存在简单关系，则经常无需通过执行具体的物理过程来得到输出。一个例子是行星未来位置的计算。如果关系不简单（程序很长），则往往没有捷径，如果要知道结果就只能老老实实地执行过程。计算机科学的确可以模拟自然过程，但如果条件复

杂，过程又长，则精确仿真所需的计算会与被仿真的现象一样复杂。这类例子包括天气、经济，以及机翼在高速状态下的性能。

如果要考虑输入输出之间的所有可能关系，复杂的会远远多于简单的。没有人知道物理宇宙是否也是这样，但这至少说明我们也许会发现不能通过简单数学认识的现象。生命、生态和大多数社会现象可能都属于这一类。

在对规则的讨论中，大部分关注的是相同初始条件导致相同输出的确定性规则。然而，大多数物理过程和一些计算机程序遵循的规则都允许随机事件。如果有随机成分，即使规则简单，也无法预测具体的结果。当这样的过程被反复执行，每次的结果都会不一样。幸运的是，在许多这类情形中，可以通过统计平均来预测，我们可以称之为总体性预测。雪花就是很好的例子。预测下一个雪花的细节是不可能的，但雪花的特征在总体上却可以预测。在某种程度上生物进化也与此类似。进化过程允许某种程度的随机变化，虽然过去或未来具体的结果无法预测，大致的趋势却可以预测出来。

科学如何解释目的复杂性？

前面我们看到可以将所有事物分为两种类型：Ⅰ型是基于简单规则，Ⅱ型则需要附加的复杂规则或指令。这样区分后，可以看到地球之所以特殊完全是因为有Ⅱ型事物，包括它们的行为和影响。这样的区分突出了指令的重要性，同时也需要解释指令中的信息来源。Ⅱ型事物还有一个特点：它们可以被视为具有某种“目的性”或“适应性”。指令总是具有目的。所有指令都引导产生特定的结果，其结果可以被认为是为了某种目的或适应某种环境。丹尼尔·丹尼特将这种对无生命对象的动机认识称为“意向性立场”[16]。

他注意到将意向性赋予无感知无思维的事物往往能正确推论出它们在未来的行为并理解它们为何如此。为什么鸟有翅膀？为了飞。为什么它们要飞？为了觅食和躲避捕食者。为什么外套隔热？为了身体保暖。为什么鱼有灵活的尾巴？为了在水里游得快。所有这些例子都与对环境的某种适应有关。这样的例子数不胜数。为什么有这么多事物都是这样？答案在于这些事物从何而来的认识。换句话说，目的性来源于创造

Ⅱ型事物的过程之中。

虽然指令有产生复杂结构的惊人潜力，宇宙中大部分结构的形成还是没有依赖它们，而且其中一些还相当复杂。一个明显的例子是星系团。第3章还会给出一些例子，但没有一个可以被合理地认为是具有目的性。基于指令的结构的另一个不同之处是它们的形成。它们需要说明书，说明书不能违背自然规律，但自然规律并不提供说明书。指令不违背化学和物理定律，而是利用它们来让本来不可能的特殊结构或过程成为可能。

目的性很早以前就受到宗教和哲学的关注。它是威廉·佩利1802年论证上帝存在的基础。他认为，如果没有造物主赋予目的，目的性事物（包括人在内）怎么可能存在？对此，道金斯在《盲眼钟表匠》中对进化如何导致目的性出现给出了优雅而有力的论证 [17] 。他认为是选择导致目的性的产生。选择是进化机制的本质之一。它决定前一代的变化有哪些会留存到下一代。个体组成的群体中会累积越来越多过去选择的特性。鱼似乎是为生活在水里设计的，是因为其祖先比那些没有留下子嗣的个体更善于在水中生存和繁衍。

加里·齐科在《没有奇迹》一书中也探讨了复杂事物如何出现的问题 [18] 。他列举了很多复杂现象，从细菌进化到人类思维，基本上都有很好的“适应”。适应有两种，一种很平常，例如打碎的瓷盘。碎片能相互拼合，因为它们以一种简单的方式源自同一个整体。还有一种能让人感兴趣的，是目的性适应。这种适应不是源自某个整体。在这种情形中，是创造了某种东西来实现适应。这类例子包括抗体对病毒的适应，生物对生态系统的适应，技术创新对用户需求的适应，思想对个人愿望和需求的适应。齐科认为，对为什么复杂事物（我们说的Ⅱ型事物）总是对某种背景有意义的问题，唯一的解释是反复的选择过程。

根据定义，指令和Ⅱ型结果总是成对出现，而指令—结果对总是有某种动机。执行指令得到的产物对象满足了指令的目的，指令则是依某种需求、愿望或可能给出。第4章讨论了螺丝刀的例子。螺丝刀有用是因为它们适应螺丝头和人手操作。螺丝刀与岩石之类的Ⅰ型事物的区别不在于是不是可能有用，而在于如果没有指令，螺丝刀的制造实现是基本不可能的。在螺丝刀的例子中，创造指令是因为人们想拧螺丝。鱼适应水是为了生存和繁殖。在这个例子中的动机是延续生命。

指令不是碰巧产生，它们的存在需要解释。在科学研究中，说某人（或神）想出了它们是不能被接受的。指令包含了特殊的信息，如果想对世界有深刻的科学认识，就必须研究这种信息的本质和来源。目前所知的唯一能产生新指令的过程就是进化计算。这里说的进化并不仅仅限于生命的进化，而是包含了一系列通过特定计算策略积累有用信息的活动。所有的进化过程，无论是否涉及生命，其背后都有一个不断循环的计算过程，这个过程能够从细微并且通常是随机的变化中提取目的性信息。这种计算的原理在于，基于现有的信息体产生细微的变化，然后根据一致的标准判断并保留能带来好处的变化。

前面我说过，Ⅱ型事物的复杂性直接源自其指令的复杂性。蛋奶酥比煎蛋复杂恰恰是因为制作所需的步骤更多，其步骤由食谱给出。总的来说，指令集越复杂，产物就可能越复杂。指令引导过程，而过程越复杂，最终的输出通常也会表现出更大的复杂性。

多步骤过程有天然的低概率性，而引导过程的指令集越长，概率就越低。与此相应的一个观察是，如果有合适的指令和执行指令的手段，产生的结构的复杂程度（和低概率性）会超乎想象。例如现代计算机，要制造这种机器，需要采矿、冶炼，制造半导体以及各种特性的塑料，再用这些材料制造出复杂的芯片和电线、电源、风扇等结构，然后还要精确地组装这些元器件。指令指导这一切：各个元器件的准确成分和设计、元器件的位置、安装顺序以及元器件的连接。只有化学和物理定律不用说明。没有这些指令，就根本不可能有电子计算机。遵循了指令，就自然会产生预设的最终结构。所有的现代技术和制造以及所有重要的人类文化创造，包括音乐、建筑、科学，甚至宗教组织，背后都是基于“指令”的原理，生物也是基于同样的原理。

在进化系统中，变化以代为单位发生，目的性通过累积选择过程一点一点慢慢地建立。例如，鱼生活在水中，生长发育机制根据DNA编码的信息形成有灵活尾部的流线型身体。编码在DNA中的信息具有让鱼在水中生活和高效移动的目的性。但并不是一开始就是如此。鱼类的祖先曾经非常小，也不是依靠肌肉运动；在水中推进是依靠纤毛，纤毛是细胞上的凸起，运动起来类似桨。这些生物也有DNA，但它们的DNA没有编码鱼的流线型身体、尾巴或肌肉的信息。这类信息之所以出现在鱼的DNA中，是因为DNA在过去的变化如果有利于鱼类适应环境就被保留了下来。原型鱼吃小东西，因此体形大都有进化优势。发生让一些后

代体形更大的突变后，更大的个体繁盛起来并因此拥有了更多的后代。但大都带来其他问题。体形大阻力也大，纤毛不再是高效的运动方式。随着时间推移，划水逐渐取代了纤毛运动，从而允许更大的体形。在自然选择的残酷压力下，游泳变得越来越高效。最终，纤毛运动被从鱼的谱系中彻底抛弃。

效率可以作为对适应性的一种度量，经常会带来更高的繁殖成功率。一旦DNA的随机变化碰巧导致了更高的效率，这种变化就会在DNA的结构中保留下来。通过反复的试错和选择过程，流线型身体、尾巴和肌肉的目的性信息慢慢累积下来，并在鱼类的DNA中逐渐完善。这个过程从没有编码如何高效游泳的信息的DNA开始，最终累积了许多相关的信息。

从计算的角度看，目的性的科学解释可以从创造指令的方式中找到。反过来，这种解释也意味着，不需要指令的事物没有目的性。它们也许会有某种用途，但它们不是为了某事而被创造出来。而指令是有目的性的，这是它们的根本定义和它们的形成方式决定的，因此它们所描述的事物也是如此。

指令的来源是进化计算，我称之为复杂引擎，在对此进行深入探讨之前，我们需要进一步探讨计算的本质以及计算、物理过程和结构形成之间的关联。

第2章　计算

什么是计算？

当你在早上启动汽车，发动机会“决定”多少汽油和空气进入气缸；随着发动机升温，油气比例发生变化。踩下油门，比例再次改变。这是怎么实现的？当你听音乐或上网时，所用的设备又是怎样的原理？所有这些事情的背后都是计算机在起作用。今天已经很难摆脱计算机的影响。计算机计算账单，无需胶卷生成照片，追踪你的喜好，预测天气。计算机是操作信息的机器，计算机科学则是研究这种活动的局限和潜力的学科。这一章我们会探讨计算的概念。我希望你能认识到计算不仅仅是计算机中发生的事情，许多物理活动都与计算密不可分。

计算机科学的基础是信息。信息可以被编码为符号，通常是向某人或某物传递意义的序列。第1章介绍了信息的一种度量：香农信息。这种度量对于一些目的很有用，比如确定存储一些你感兴趣的信息需要多少物理空间，或将其远距离传送需要多大带宽。这一章我还会介绍另一种度量——算法信息，它对计算机理论也特别有用。

理解计算的概念的一个可行途径是将所有计算都视为一个过程，过程以编码模式（通常是符号序列）为输入，经过一系列变化，最终产生输出。输出可能有输入所不具备的用途。这个大致的思想表示在图2.1中。

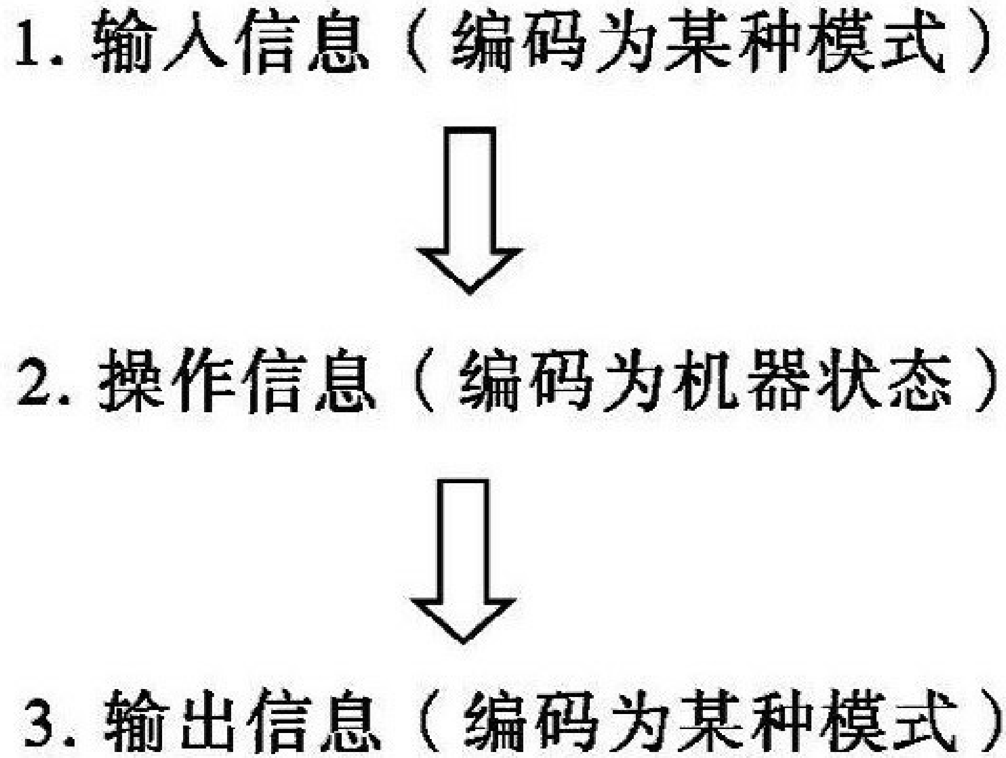


图2.1　计算的3个部分

这引出了几个问题：操作信息是什么意思？计算中会损失信息吗？新的信息来自哪里？计算机要能“操作信息”的必要条件是什么？要回答这些问题，我们需要一些背景知识。

我们从物理中“系统”的思想开始。系统是具有边界的某种事物。虚空不是系统；数学上的点或几何平面这类抽象思想也不是。宇宙可能是也可能不是系统，取决于其是否有边界。所有比有限宇宙小比亚原子粒子大的事物都是系统，也是更大系统的一部分。系统通常由多个部分组成，各部分之间可以有不同的组合方式。我们将组合方式称为系统的“状态”。在第1章我们看到，状态如果可以相互区分，就能够携带信息。至于信息是否有用则是另一个问题，取决于是否存在另一个系统通过与原系统的特定状态互动产生特定的变化。以开车为例，你自己是一个系统，车是另一个系统。两个系统互动。坐在驾驶席上脚踩油门踏板的你处于某个状态。脚踩下去你就变成另一个状态；这个新的状态导致汽车系统转变成一个新的更快的状态。你的某些状态就对汽车系统有意义；还有一些比如闭上一只眼睛或挥动左手则没有。在物理的世界里，

意义来自系统状态的互动。如果没有互动，就没有意义。要具有意义，一个系统的特定状态必须对另外的系统有特定的影响。

在信息科学中，意义经常与表示的概念联系到一起。计算的背后有一条原理是一个状态可以表示为另一个状态。这里的状态不必是在同一个系统。举两个例子，你在键盘上敲的字可以表示为计算机里的电压和电流；人类艺术家演奏的音乐可以表示为DVD上的银点样式。当某种表示影响到了你，它就对你有意义。

计算机程序是由能导致计算机内部发生特定的状态序列并产生输出的命令组成。在计算机科学中，完成某件事的命令序列被称为算法。在其他场景中可能被称为食谱或指令。指令和算法表示完成任务的方法。大部分计算机算法都接受数据输入并产生针对某种场合的输出。数据和算法都包含信息。数据可以视为外部世界某方面的表示。这样来看，计算就是将对方法的表示（算法）作用于对世界某方面的表示（数据）产生内部表示（机器状态）的序列，并得到最终的表示（输出）[1]。无论哪种表示，都是某个系统的物理状态，也都可以理解为信息。这可能让人感觉很抽象，但让人吃惊的是，这可以用物理设备实现，并最终解释为什么信息在许多场合中对我们如此重要。计算机就是可以呈现许多内部状态的设备。在计算机中，内部状态与输入状态互动产生新的内部状态。最终状态就是输出。

计算的现代概念已经发展了几十年。在有计算机之前，人们通过演算解数学题，数学家和哲学家则希望能用机器模拟这个过程。20世纪30年代后期，年轻的英国数学家阿兰·图灵建立了现代计算机科学的数学基础。数年后他领导一个英国团队成功破解了德军密码。图灵对第二次世界大战的贡献很重要但鲜为人知，以至于在战后很难找到工作。直到他在1954年自杀后，他对战争的贡献才被广泛认可。图灵最大的科学成就源自对人如何解算术题的仔细分析。例如用58432除以83。大部分人都可以不借助计算器通过一些简单步骤演算出结果。小学就教了这个。表2.1列出了20个步骤，只要照着做就能得到正确答案。就连我这样在小学时讨厌长除运算的人都记得这些冗长的步骤。如果你不想一行行读这张表，就不用读它。

表**2.1**　**58432÷83**的**20**个简单分解步骤

1. 将数字 58 432 写在纸上。

2. 将 83 依次与 5、58 和 584 比较，直到找到比 83 大的数。

3. 由于 83 小于 584，猜测一个最大的小于 10 的整数，与 83 相乘的结果小于 584。

4. 在 58 432 中的数字 4 上面写下这个数。

5. 将刚才写下的数与 83 相乘。

6. 将乘积写在（58 432 中的）584 下面并相减。

7. 如果第 6 步得到的差为负，擦掉第 4 步写的数，减去 1，然后重新进行第 4 步。

8. 如果第 6 步得到的差为正并且大于或等于 83，擦掉第 4 步写的数，加上 1，然后重新进行第 4 步。

9. 如果第 6 步得到的差为正并且小于 83，将（58432 中的）3 附在其后面。

10. 将得到的数与 83 比较。

11. 如果大于 83，猜测一个最大的小于 10 的整数，与 83 相乘的结果小于这个数。

12. 将这个数写在 58 432 的 3 上面。

13. 如果第 10 步是小于 83，在 58 432 的 3 上面写 0，然后将（58 432 的）2 附在第
9 步得到的数的后面（因为第 9 步的相减得到的数小于 83，我们略过得到其他结果后的不同步骤）。

14. 如果第 13 步得到的数大于或等于 83，猜测一个最大的小于 10 的整数，与 83 相乘的结果小于或等于这个数。

15. 将猜的数写在 58 432 的 2 上面。

16. 将第 15 步写下的数与 83 相乘，将结果写在第 13 步得出的数下面。

17. 相减。

18. 如果 17 步得到的数为负，返回第 14 步，将猜的数减 1。

19. 如果 17 步得到的数为正并且大于或等于 83，返回第 14 步，将猜的数加 1。

20. 如果 17 步得到的数为 0，计算过程结束，58 432 上面写的就是正确答案（704）。

任何人只要遵循这些步骤就能得到正确答案。他们无需“理解”他们在做什么，也无需具有在第3、第11和第14步一猜即中的天赋。他们只需知道判断一个数比另一个数大还是小，以及做乘法和减法。如果他们不知道乘法，也可以将乘法还原成重复相加（步骤会增加）。图灵认为任何计算，无论多复杂，都能通过顺序执行极为简单的步骤解决——甚至比表2.1中的步骤还要简单。由于这些步骤很简单，不需要什么聪明才智，因此可以设计一台（不会思维的）机器来执行这些步骤。然后他展示了如何设计一台这样的机器。为了纪念他，他提出的这个想法现在被称为图灵机，所有计算机专业的学生都会学。现代计算机的设计已经有了很大变化，但仍然是基于图灵的思想。原则上，图灵机可以通过顺序执行适当的无需思维的简单步骤来解决任何可计算的问题。

第一台实用的计算机在第二次世界大战期间被建造，通过反复解数值问题计算炮弹弹道，通过系统地对数百万种可能性进行测试破译军用密码。无论具体的设计和执行计算的细节如何，所有计算机执行的都是一长串极为简单的步骤。它们的优势在于能高效准确地执行这些步骤。

表2.1给出的20个步骤就是一个算法。算法的一个显著特征是它们的逻辑结构超越了对它们进行编码的物理系统。这意味着算法的结果独立于执行算法的机器。用丹尼特的话说，无论是“在纸上、在羊皮上、在计算机里……还是在天上”，长除问题的答案都是一样的 [2] 。这是因为算法可以用不同的物理媒介甚至不同的语言表示。其超越性来自符号序列可以被复制、重排、转换成其他符号形式，而不会失去信息本身的意义或逻辑结构。算法的一个重要特征是必须有结果。一系列命令如果不能产生输出就不是算法，也没有意义。

计算机接受的程序（算法）是用由顺序排列的符号组成的语言写出来的。输入符号的顺序驱使机器做特定的事情。例如，在许多计算机语言中，命令“x=3”会使得计算机将数字3存储在标号为x的存储器中，随后的命令“print　x”就会使得计算机将x的内容（例如3）送到打印机或屏幕上。符号的顺序决定一切；“rpint　x”或“x　print”都无法让数字3出现在屏幕或打印纸上。

计算的概念不限于编码为符号序列的信息。2维和3维样式也可以被视为信息。不过，大家包括计算机专家在内都喜欢1维，因此几乎所有人造的计算机都是基于对顺序符号序列的操作。

操作信息是什么意思？

在计算的过程中，计算机产生一系列内部状态，每个状态都由之前的机器状态和新的输入决定。状态链条从输入数据、算法（规则）和机器的初始状态开始。所有这些，输入、规则、内部状态和输出，都编码信息。由于信息最终都位于状态中，计算机就可以视为一个状态（输入）与另一个状态（当前机器状态）互动产生最终状态（输出）的设备。所有状态都依约定表示什么。最简单的计算只涉及两个状态，初始状态和最终状态。

计算机通过历经由算法和输入数据决定的机器状态序列执行计算过程。字处理软件体现了这个思想。我逐次键入字母和空格（输入），然后看见它们成行成页出现在屏幕上（输出）。每一次键盘敲击与机器当前状态的互动导致形成新的内存和处理器状态，继而又导致屏幕上的特定状态。改变输入就会导致不同的内部状态产生不同的输出，但过程始

终完全由输入、程序和机器的设计决定。

计算的广义概念涵盖了范围广泛的活动，可以分为3个宽泛的类别：可逆的、确定性的和随机的。每个类别都各有特点。你的计算机是确定性的并且是不可逆的。因为是确定性的，每次你运行同样的程序都会得到同样的输出。可逆计算指的是既能往前也能往后运行的计算。这意味着可以根据输出得到输入。前面我们考虑了除法问题58432÷83。解决这个问题得到答案704的过程是不可逆的。如果只知道数字704，无论是计算机还是你都无法重构出具体的问题。要让计算可逆，答案就必须包含重构所需的足够信息。比如“704是某数除以83的答案”就能通过适当的算法重构出原来的问题。

随机计算每次运行都会产生不同的输出。至少要有一条规则具有随机行为。随机计算既不可逆也无法重复。但它们能产生你想要的一切，只要你愿意运行计算足够长时间。现代数字计算机是确定性的，设计者花费大量心力就是为了确保这一点。如果一台计算机对相同的输入每次都得出不同的输出，则对大部分应用都不会很有用。

数字计算机之所以可以实现，是因为可以制造出能基于是/否、通/断或电平高/低的顺序组合执行逻辑运算的物理设备。在计算机中，从输入和初始的设备状态开始，顺序产生出设备状态（电平高低或开关通断的组合）的链条。将高电平作为1，低电平作为0（反之也行），就能用状态表示数字，计算机就能通过序列的高低电平逻辑操作实现数学运算。每一步操作都是图灵定义的作为所有计算的基础的简单步骤，引用名为逻辑门的电路实现。

图2.2展示了两种逻辑门以及由它们组成的一个简单逻辑电路。电路用高低电平分别表示1和0，可以执行简单的逻辑运算。上面的符号代表XOR（异或）门，中间的是AND（与）门。两者都通过输入线（图中标为A和B）接收高低电平。AND门的两个输入如果都是1（高电平），输出就为1，否则就输出0（低电平）。XOR门的两个输入如果有且只有一个为1，输出就为1，如果两个输入都为1或都为0，输出就为0。如果将这两种逻辑门像下图那样连接起来，电路就能正确执行两个1和0的各种组合（0+0、0+1、1+0和1+1）的加法。如果输入A是1，B是0，上面的输出就是1，下面的是0。十进制数1的二进制表示是01（参见表2.2）。如果两个输入都为1，输出就为10（十进制数2）。这个简单

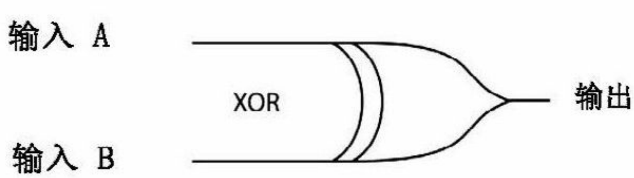
的加法器不能输出11（十进制数3）；它只能加到2。这个支配加法器功能的“规则”隐含在电路的物理结构中。更复杂的逻辑电路可以执行多位数的加减乘除[3]。

表**2.2**　二进制数和十进制数的对应关系

二进制	十进制
00	0
01	1
10	2
11	3

计算机也可以建造成不需要通过硬连线而是以可编程的方式实现逻辑运算。也就是说图2.2中的XOR门和AND门不用像图中那样固定，而是可以与其他门用各种方式组合起来完成不同的逻辑运算。在这样的机器中，逻辑由程序决定。

可以将逻辑和要操作的信息都作为输入，此时一般将输入分为两部分：程序和数据。程序可以视为由机器内建的（基础性的）规则集组成的规则。计算过程中根据程序规则操作数据。现代数字计算机通过编程基本可以实现任何逻辑运算。具有这种特性的机器称为通用机。其他计算机能够执行的计算，通用计算机都能通过编程实现。通用性定义了很宽泛的一类操作，称为可计算数学函数，其中涵盖了几乎所有已知的数学。这就是为什么电子计算机的用途如此广泛。



A	B	输出	电平		
0	0	0	低	低	低
0	1	1	低	高	高
1	0	1	高	低	高
1	1	0	高	高	低



A	B	输出	电平		
0	0	0	低	低	低
0	1	0	低	高	低
1	0	0	高	低	低
1	1	1	高	高	高

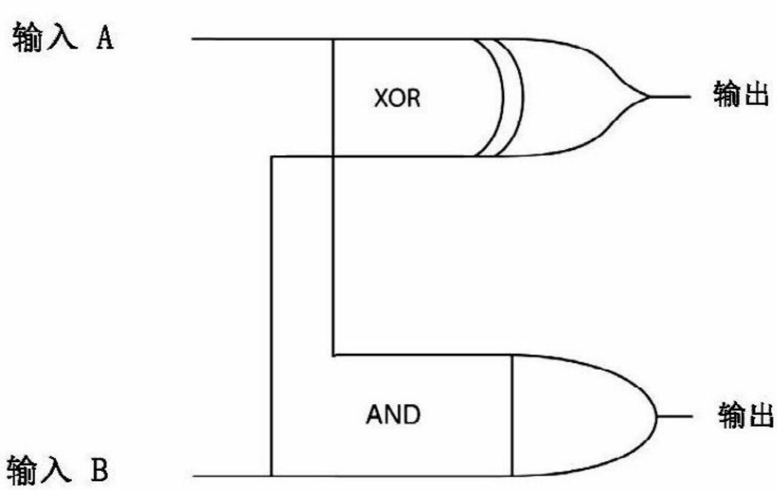


图2.2　上图：XOR门；中图：AND门。表中列出了所有可能输入对应的输出。下图：2比特加法器。输入A和B可以为1或0。两个输出组成了由输入决定的二进制数（00、10或01）

计算机科学如何定义和度量信息？

对于确定性计算，输出中呈现的信息必然也以某种方式呈现在输入中——信息必须来自某处。计算机不会产生信息；它们被重排、计算、复制和删除，但不会无中生有。古谚“太阳底下无新事”道出了确定性计算的精髓。第1章曾说过，对字符串信息量最简单的度量是长度。不过这个简单的度量没有反映出确定性计算不会产生信息的事实。显然许多计算产生的输出要比输入长。

有一种被计算机学家广泛使用的度量，它来自一个认识，就是有许多方式都可以计算出特定的输出。实际上，有无穷多种方式。如果你不信，想一想有多少算术题的答案为8。4+4=8，2+6=8，12-4=8，440-432=8，还有很多，要多少有多少。但是对于任何给定的输出（答案），使用指定的语言，至少有一个最短的输入，即没有更短的输入能产生同样的输出。可能有几个最短输入，都有同样的长度，但总存在一个最短长度，没有长度比它低的输入能产生给定的输出。

为了让你信服，考虑上面的例子，答案为8的所有算术题。显然，4+4（3个符号）比12-4（4个符号）或440-432（7个符号）短。有10个3符号的算术题的输出为数字8：0+8，1+7，2+6，3+5，4+4，8-0，9-1，1×8，2×4和8÷1。它们的长度一样，并且没有比这更短的。如果用二进制编码，这10个表达式的长度不一样，但上面的原理仍然成立：对任何输入，给定计算机和语言，存在某个输入长度，低于这个长度没有输入能产生给定的输出。

20世纪60年代末，苏联数学家柯尔莫哥洛夫将这个认识形式化，并据此给出了不同于香农度量的信息定义。这个度量被称为算法复杂性或柯尔莫哥洛夫复杂性，其基础是通用计算机。前面说过通用计算机是可编程的，因此，它能执行其他计算机能执行的任何计算。你的电脑就是一台通用计算机。功能简单的计算器则不是。

大体上，一个二进制串（0和1组成的串）的算法信息量就是输出为这个串的特定通用计算机的最短输入的长度。这个输入的长度依你选择的计算机和语言的不同而有所不同，但柯尔莫哥洛夫从数学上证明了这个长度的差异不会超过一个常数，这个常数只取决于选择的电脑，与输出无关。这意味着对于很长的输出（远远大于常数），无论采用什么通用计算机（和语言），得出的算法信息量基本是一样的。如果我们选择的机器接受自然语言命令，输入程序“重复1000次01”产生的输出由2000

个交替的0和1组成（2000比特长）。这个程序如果直接转化成二进制大约需要50个0和1（50比特）。因此我们可以确定这个2000比特的输出的算法信息量不超过50比特。如果选择其他语言，上限可能可以降低，但绝不会低于15比特，因为重复的串（01）、命令“重复”和数字1000都对算法信息量有贡献。

算法信息与香农信息的共通之处是确定性计算不能创造信息。输出的信息量可以比输入少，但绝不会多。因此算法信息符合我们的直觉，即信息无法无中生有。

如果用一个对象的长度减去其算法信息，差值就是信息的“冗余”或“可压缩性”。这表明一个巧妙的对象表示可以比对象本身短多少仍可以通过适当的确定性计算恢复出对象。压缩文档比原文档短，但可以恢复出原来的0和1的排列。

随机性的概念很难给出精确的数学定义，算法信息提供了一个有力的定义。如果算法信息量和符号序列的长度相等（准确说是差异小于由使用的计算机决定的一个小常数），则序列的结构被认为是算法随机的。这意味着任何确定性计算如果要将这样的序列压缩得更短，都会不可逆地损失一些信息。具有这种属性的序列没有冗余，被认为是随机的。

这种随机性定义揭示出算法信息一个有趣的特性。当比较两个相同长度的序列时，最接近随机的那一个具有最高的信息量。当我们考虑复杂事物时，通常不会认为它们是随机的，而是具有复杂的特定结构。为了认识随机性与复杂性的关系，想一想重现一个特定的序列需要准确说明每一个符号。序列0101010101……（重复1000次01）有许多重复成分。因此可以写一个很短的程序输出这个序列。随机事物缺乏重复性，因此要输出一个特定的随机序列的程序就必须以某种方式包含序列本身。复杂事物的特点是有许多细节，要产生所有这些细节需要很长的程序，因此复杂的事物其算法信息度量有可能也高，但并不总是如此。

简短程序能产生大的复杂性吗？

对逻辑运算的数学分析揭示了确定性计算能做什么和不能做什么的基本限制。以下3条原理很重要：

原理1：简单输入和简单变换规则能（但并不必然）导致输出具有大的复杂性。

原理2：如果简单变换规则作用于简单输入产生大的复杂性，则变换需要很多计算步骤。

原理3：许多复杂对象的形成需要复杂的程序。这类对象无法通过简单规则作用于简单输入产生，无论多少计算步骤也不行。

沃尔弗拉姆深入研究了原理1，并在《新科学》一书中展示了他的许多发现 [4] 。他将NKS（新科学）系统定义为通过简单规则和输入产生大的复杂性的逻辑系统，并认为我们在物理世界中观察到的大部分复杂性都是原理1的体现。

图2.3到图2.6展示了4个NKS系统。这些例子被称为元胞自动机（CA），其状态是根据一些简单的规则随时间逐步变化。有许多NKS系统并不是基于元胞自动机，但这种计算能很好地说明这个原理。元胞自动机是元胞（不是生物学意义的细胞）组成的1维、2维或3维阵列，其状态定期变化。在图2.3中，规则集由左边图示的8条规则组成。每条规则由1个大格子中的4个小格子描述。左边显示的规则相对右边的计算区域旋转了90度。第1条规则的意思是，如果某一步相邻的3个元胞都是白的，则中间元胞下一步是白的。第2条规则的意思是，如果3个相邻元胞最右边的是黑的，则中间元胞下一步会变成黑的，以此类推。根据这8条规则，显然只要3个元胞有1个是黑的，中间元胞下一步就是黑的，只有3个元胞都是白的，下一步的中间元胞才是白的。

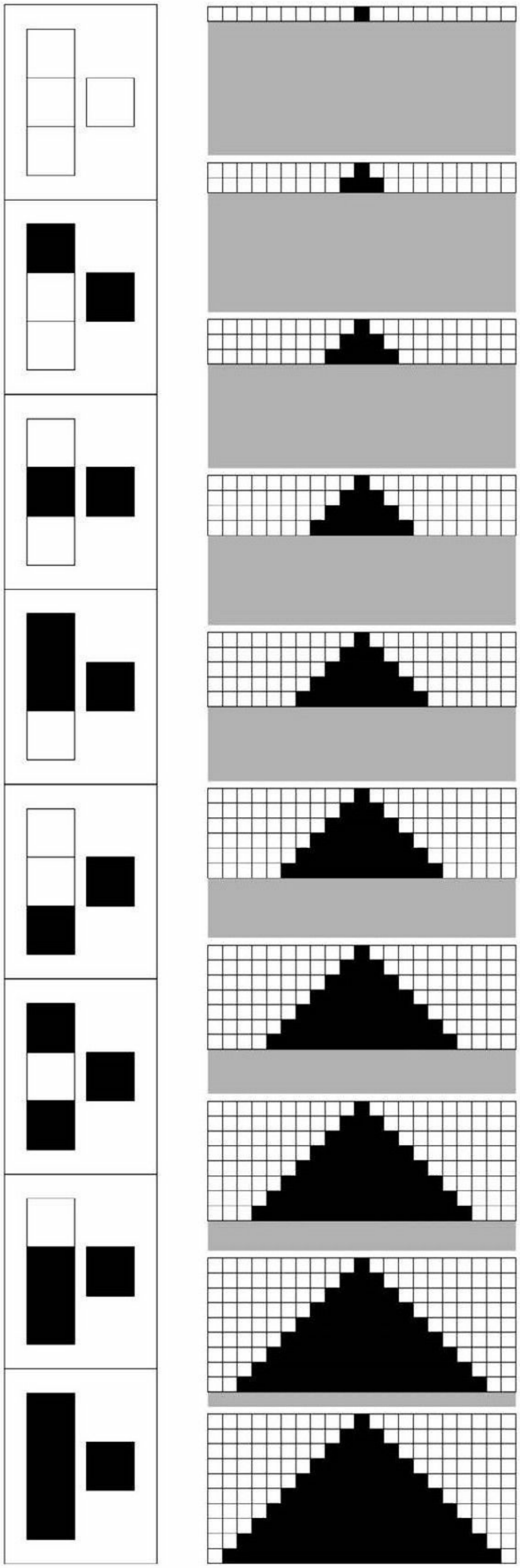


图2.3　应用规则集254（左侧）的元胞自动机会产生一个一直生长的黑三角形（右侧）。沃尔弗拉姆研究公司版权所有（参见注释4）

从第1行元胞开始用规则集执行计算。你可以从任意的黑白图样开始。图2.3右边的初始行只有1个黑元胞，其余都是白的（第1步）。然后应用规则集254。产生的行有3个黑元胞，其余都是白的（第2步）。

继续执行，会基于第2行产生出第3行元胞图样（第3步）。这一行有5个黑元胞。计算可以一直继续下去。可以看出，从这个简单的输入数据开始（第一行），应用这个简单的规则，结果很简单，就是一个生长的黑三角形。规则集254无法产生出比输入行明显更复杂的图样。图2.4展示了用不同规则集实现的一个更有趣的元胞自动机。这个例子也是从只有一个黑元胞，其余全白的初始行开始。产生的图样被称为塞平斯基填充，这是数学中一个经典的分形图形。继续更新，会产生越来越大的三角形以及各种大小的小三角形。

如果元胞限于2种颜色（黑和白），并且每个元胞的颜色由上面的3个相邻元胞的颜色决定（如图2.3和2.4中的例子），则覆盖所有可能需要8条规则，因此总共有$2^8=256$种规则集。这些规则集大部分都只能产生简单图样，或是在第1行之后根本没有图样。少部分会产生更复杂的图样。

图2.5展示了规则110作用于只有单个黑元胞的输入行的结果，产生的图样规则与不规则并存。令人吃惊的是，规则110是通用性的。意思是说通过让规则110一遍又一遍作用于适当的输入行可以构造一台通用计算机。这个理论结果证明了通用计算逻辑上可以很简单，简单规则也能产生大的复杂性。

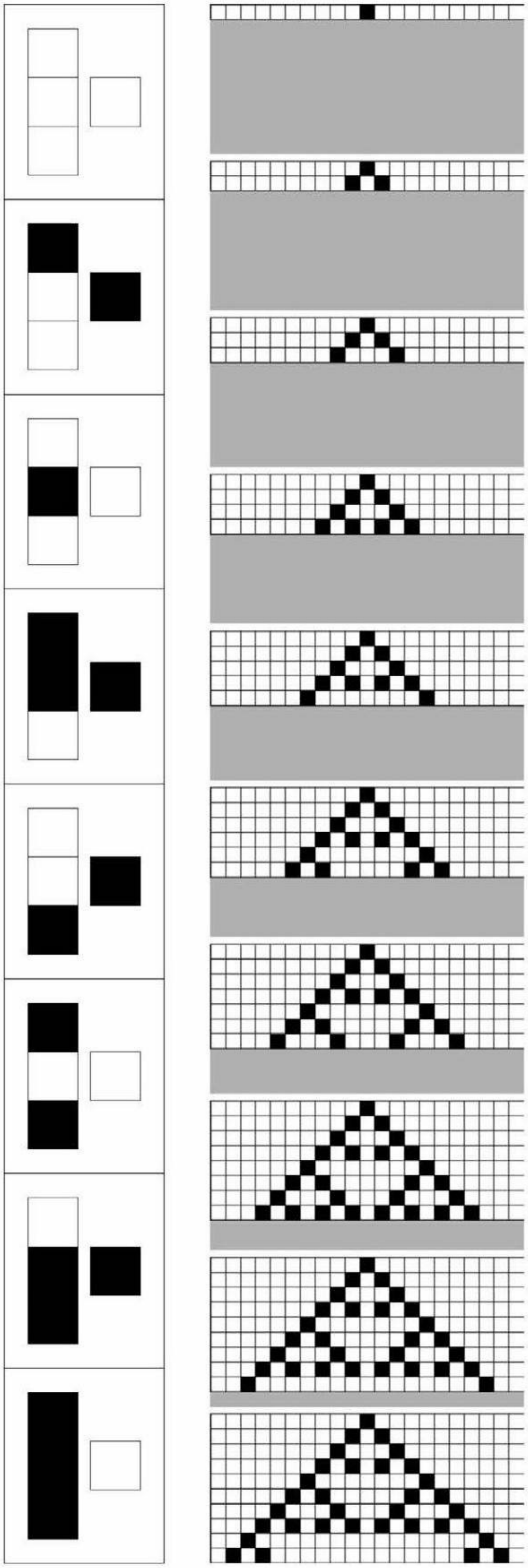


图2.4　实现规则集90（左侧）的前10步的元胞自动机。这个图样被称为塞平斯基填充。沃尔弗拉姆研究公司版权所有（参见注释4）

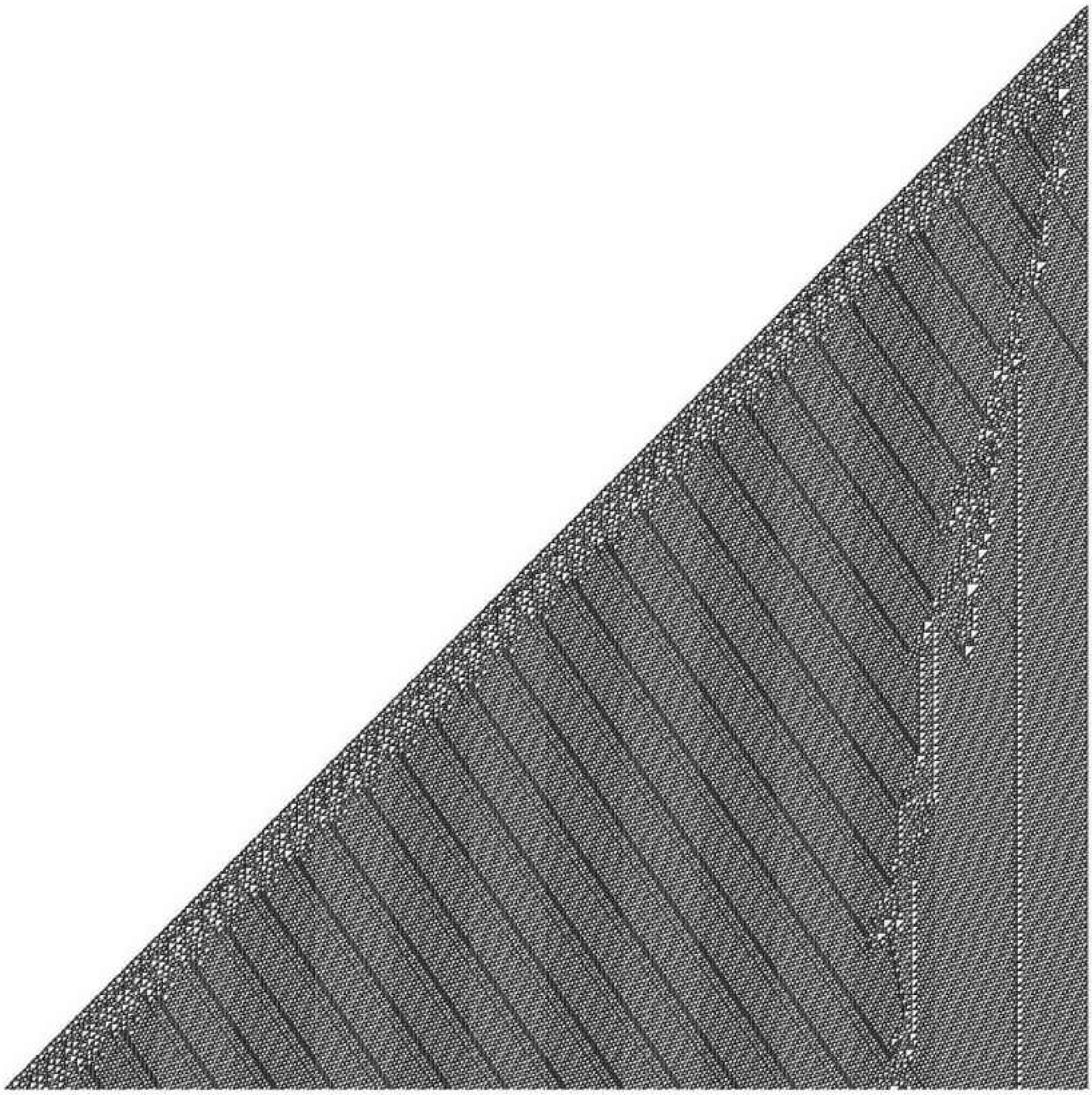


图2.5　实现规则集110的元胞机前1000步产生的图样的左半边。沃尔弗拉姆研究公司版权所有（参见注释4）

如果将元胞的颜色增加到3种（白、黑和灰），全部规则的数量从8条增加到27条，规则集的数量则增加到700多万种；但是同样，只有小部分规则会产生复杂的图样，大部分还是只能产生简单输出。

稍微改动一下方案，让元胞的颜色由上面3个相邻元胞的灰度均值

决定。3个元胞的颜色均值有7级灰度，因此需要7条规则组成规则集：白（3白）、浅白灰（2白+灰）、白灰（白+2灰或2白+黑）、灰（3灰或白+灰+黑）、黑灰（2灰+黑或白+2黑）、深黑灰（灰+2黑）和黑（3黑）。这样总共有3^7 =2187种规则集。同样，使用这种规则集的元胞机大部分都只能产生简单输出，但少数会产生相当有创意的图样。图2.6展示了规则集1635作用于只有一个黑元胞，其余全白的初始行产生的图样。

这4个简单的元胞机例子体现了简单规则反复作用于极为简单输入（这里是只有1个黑元胞，其余全白）会发生什么事情。显然对这一节标题中提出的问题的答案是，一些简单规则作用于简单输入能产生显著的输出复杂性。计算机学家设计了各种系统来测试将简单规则作用于简单输入，结论都是简单规则和输入的结合能产生显著的复杂性，但并不普遍，即便是在那些能产生复杂性的系统中，大部分规则集都只能产生简单结果[5]。

图2.6说明了我们讨论过的几个信息度量并不符合我们日常的复杂性观念。所有人都会同意图中的图样很复杂。图2.6靠近底部的一行元胞似乎编码了很多信息，因为很复杂很长。但我们知道它的算法和香农信息度量很低，因为产生它的规则集和输入数据（1个黑元胞）很简单。因此，我们讨论过的这些信息度量并不是人类观察者所认识的复杂性的可靠指标。

NKS系统的行为通常用计算机进行探索，但我们看到的这4个例子也可以通过在桌上排列黑白灰卡片完成。后面我们会看到在试管中也有类似行为的系统。可以将其视为是通过自组装名为分子砖的小单元执行计算。这种计算的输出是物理结构。

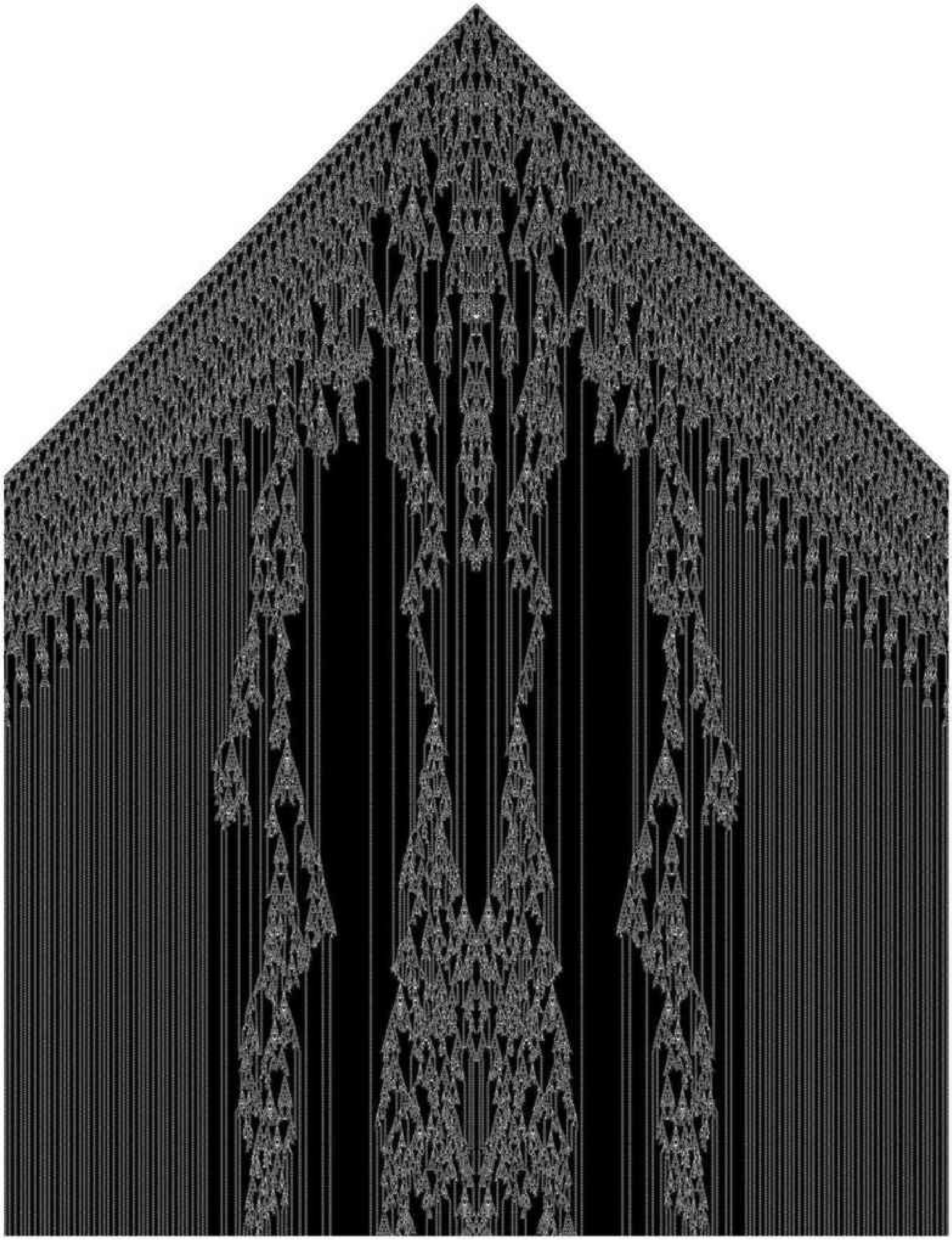


图2.6　规则集1635前3000步产生的图样，采用的是3种颜色和均值规则。沃尔弗拉姆研究公司版权所有（参见注释4）

物理结构能够计算吗？

在现代的科学世界观中，是很少的一些相对较简单的规则作用于简单的初始条件，经过漫长时间的演化，产生出大量的复杂事物。物理系统同计算系统一样受限于规则，但这里的规则是化学和物理定律，而不是来自人的设计。如果物理定律没有逻辑可循，数学就无法刻画自然，也不会有物理学这门科学。规则在计算和物理系统中的同等地位引发了长期的争议，人们怀疑是不是可以将宇宙视为一台计算机。这一节并不想尝试回答这个问题，而是阐释为什么不可能在物理过程和计算之间划分明确的界线。

加州理工学院的埃里克·温弗里与其同事创造了行为类似于1维元胞机的物理系统[6]。这种系统是基于名为“分子砖”的化学结构。砖块是扁平结构，有两条边可以与其他砖块结合。通过适当的设计，砖块可以“自组装”成预先设定的结构。

扁平砖块只能通过边结合形成1维或2维结构，形成3维结构则需要3维砖块。图2.7a上面一行展示了4个扁平砖块[7]。每个砖块两端都有A型或B型结合点。假设有2条规则——A只能结合B并且A不能与倒置的B结合（B也不能与倒置的A结合），那这4个砖块结合的方式只有一种，产生的结构如下一行所示。

图2.7b展示了端部有4个结合点的3个砖块。这个例子中结合点有4种类型：A、B、C和D。基本规则是A与B结合，C与D结合。同样不允许倒置的结合。这些砖块上每个标有字母的结合点都有特定的形状。结合点的特定组合可以形成交错的结构，就像下面的3个砖块那样。

在图2.7c中用了给出的4种砖块建造一个生长的结构。只有互补的形状才能结合。如果规则要求只有一端的两个结合点都能匹配的砖块才能结合，那图2.7 c上面的砖块就不能生成任何结构，因为这些砖块相互之间都无法结合。

为了解决这个问题，必须从一个模板或种子行开始组装过程。图2.7c下面展示的生长结构从模板行开始，模板中间部分有一个灰色的结合点。Ⅰ型砖块可以在模板的任意位置结合，只有那个用星号标记的灰色位置除外。这个位置只有Ⅲ型砖块可以结合。在模板上组装的第1行

砖块上面又可以形成第2行。第2行大部分也是由Ⅰ型砖块组成，但是在第1行唯一的Ⅲ型砖块上面相邻的是Ⅲ型和Ⅵ型砖块。第2行上面又可以组装第3行，可以一直进行下去。每一行都是由下面那一行的样式决定。如果将Ⅰ型和Ⅱ型砖块填成灰色（如图所示），将Ⅲ型和Ⅳ型填成白色，则白色砖块的生长样式为塞平斯基填充。与图2.4元胞机规则集90计算出的样式一样。模板上唯一的Ⅲ型结合点在物理拼砌系统中起到的作用与图2.3到图2.6的1维元胞机中唯一的黑色元胞一样。

图2.7　c是画出来的。图2.8中展示的塞平斯基填充则是在试管中由DNA短链砖块自发形成的微观物理结构。这些物理砖块是图2.7c中设计的砖块的化学版本。图2.8中的结构是通过在适宜的盐水中加入混合的DNA砖块和少量模板形成的。原则上，适当设计的砖块和模板组成的分子拼砌系统能生成任何想要的序列。这意味着，理论上，电子计算机能实现的任何算法也能通过适当的砖块集合和模板的组装实现。这也意味着，只要有适当的砖块集合和模板，任何可以构想的2维结构都可以通过自组装创建出来。推而广之，“物质可以通过编程从简单的成分自我组装成有用的结构”[8]。

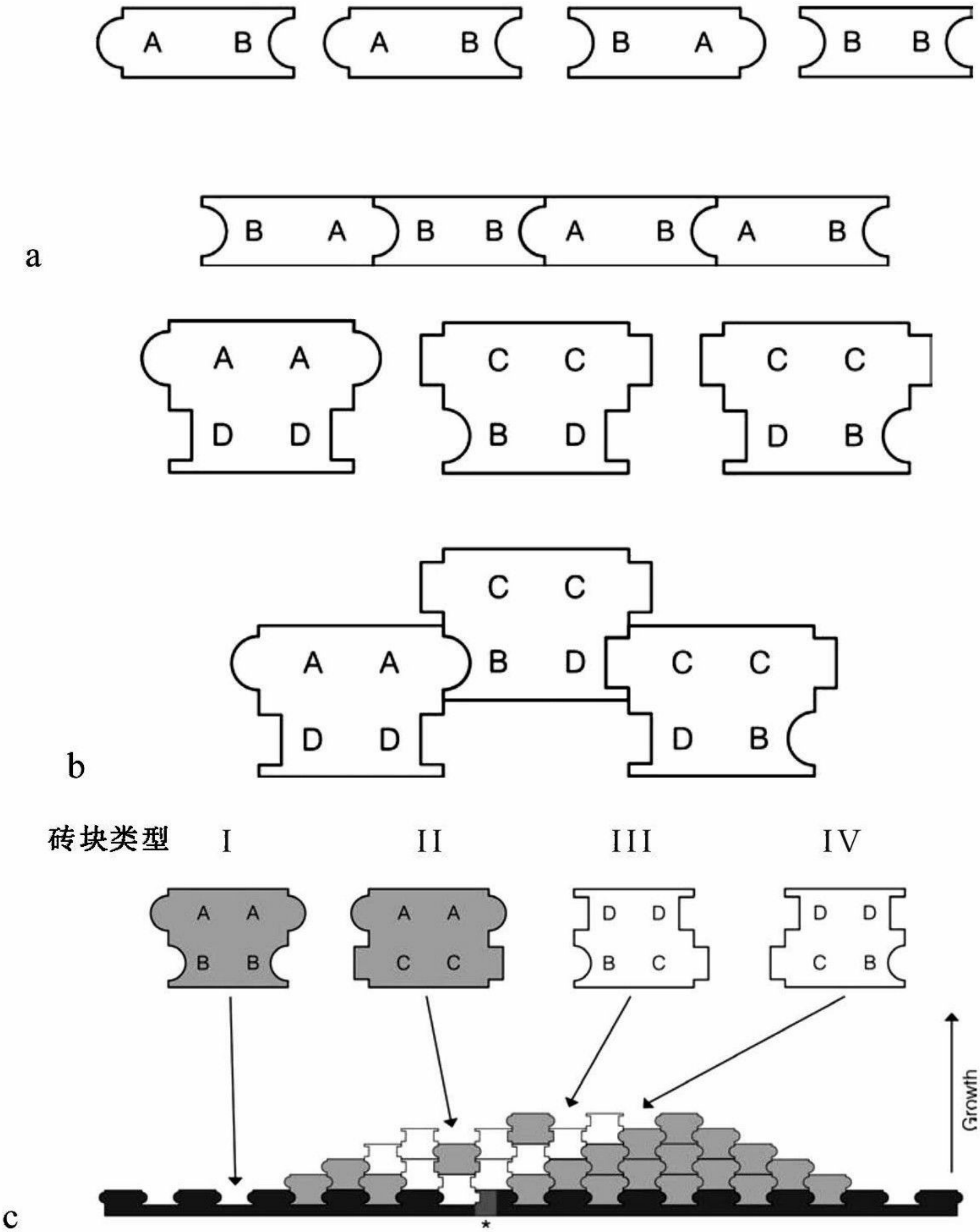


图2.7　a.两端各有1个结合点的自组装砖块图以及砖块形成的结构。b.有4个结合点的自组装砖块。这种特定的结合点设置只允许交错匹配。c.4种类型的砖块在模板（黑色部分）上的组装。在模板上的第一行，Ⅰ型砖块可以结合任意位置，除了一个位置（用星号标记）。III型砖块可以结合这个位置。接着往上的每一行都表现出新的特点（参见注释7）

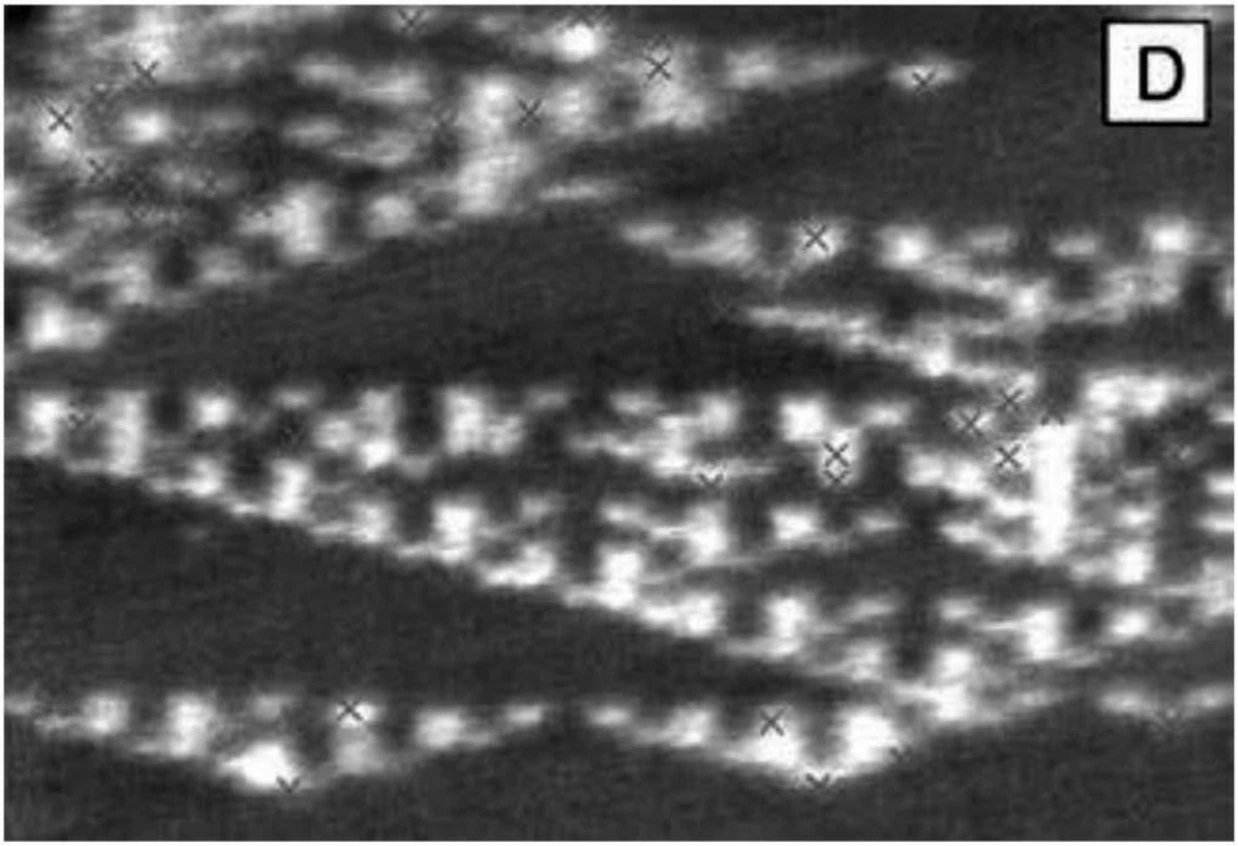


图2.8　原子力显微镜下DNA短链砖块形成的三角形图样。白色砖块被设计得比灰色砖块突出，形成颜色对比。错误标记为x（参见注释7）

埃里克·温弗里在他的博士论文中研究了砖块组装的潜在计算能力，他证明了可以设计出能执行（通用）图灵机计算的砖块集。这表明电子计算机能计算的原则上分子砖也能计算。

目前的拼砌系统由于错误率很高，实用性还很有限。分子系统不像电子计算机那样表现完美，因为分子结合有时候会出错。尤其是两个结合点必须绑定到一起砖块才能形成大的结构的规则很难保证遵守。在图2.8的实验中就出现了一系列错误。在简单的自组装系统中，错误的传播会摧毁组装的确定性，从而限制能够成功形成的结构的复杂性。不过我们知道，还是有可能通过自组装形成非常复杂的结构。毕竟生物体就是这样的结构。

错误传播的问题并不是只有分子砖组装系统才有，实际上所有形式的确定性计算都会有。你的计算机通过将错误率降到极低水平，有时候还结合纠错策略，从而避开了这个问题。纠错指的是在继续进行下一步

之前检查上一个步骤是否执行正确的机制。在分子尺度的过程中，热噪声是无法避免的。因此许多基于分子组装的计算实现都必须有纠错机制。生命系统的存在依赖于复杂结构的分子自组装，因此也具有许多类似的策略。其中许多还无法有效地在人工的试管自组装系统中实现。

物理对象的自组装的特点是通过砖块设计来物理体现计算规则。可以通过图2.7c与图2.9的对比说明这一点。图2.9中标有1和0的方块形成了交错的行。在图2.7 c中砖块基于互补形状形成行。图2.9展示了计算规则："如果下面的2个砖块有1个为1就添加1块1型砖，如果下面2个砖块都为0或都为1，就添加1块0型砖。"这就是图2.2中用逻辑门实现的"异或"（XOR）逻辑函数。塞平斯基填充就是从一个单独的1开始，通过反复应用XOR逻辑函数得到的。规则集90（图2.4）是XOR逻辑函数的另一个实现。图2.8中分子结构的建构则是在化学上重复应用同样的XOR逻辑函数。这个规则没有明说，而是隐含在砖块的设计中。不同的砖块设计能实现其他逻辑函数并形成其他结构。

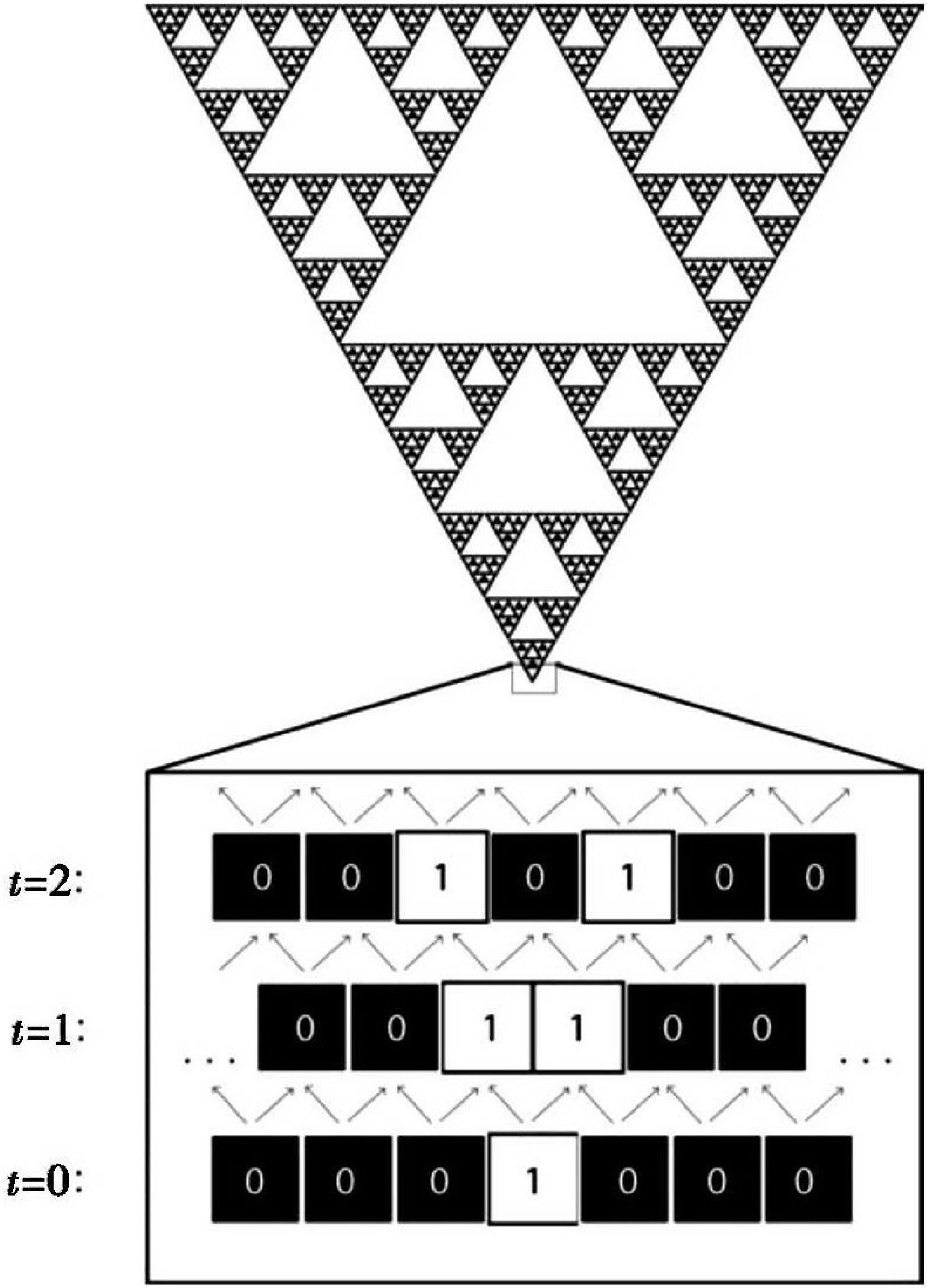


图2.9　用砖块实现塞平斯基填充的背后的逻辑。t=0，t=1和t=2表示依次产生的行，每一行的图

样都是取决于下面的行（参见注释7）

图2.9也能帮助我们理解种子序列的作用。塞平斯基三角形的生长是通过将XOR函数作用于包含一个1其余全为0的初始行。如果种子行由隔得很开的1组成，则每个1都会长出塞平斯基三角形图样。最终，三角形图样会碰到一起，图样相互干扰，变得很复杂。由0和1的复杂样式组成的初始种子序列会导致后续的0和1样式同样复杂。通过巧妙地安排种子序列，只需用XOR函数作为规则，就能计算出特定的复杂序列。但XOR规则还不足以实现通用计算。你需要规则集110（图2.5）那样特殊（并罕见）的规则。

想象一下人类活动尺度的砖块自组装，假设在拼图的每一块的边缘特定位置附上小磁铁。磁铁的强度被设置成只有两对磁铁正确对齐才能精密绑定到一起。这种特殊的拼图具有魔术般的特性，只需将图块放在平面上，所有面朝上，然后水平摇晃，让图块跟着晃动但不翻转，就能组装起来。最终，图块会以合适的角度相遇并结合到一起。慢慢地拼图就会自行组装。图2.10展示了这样的例子。

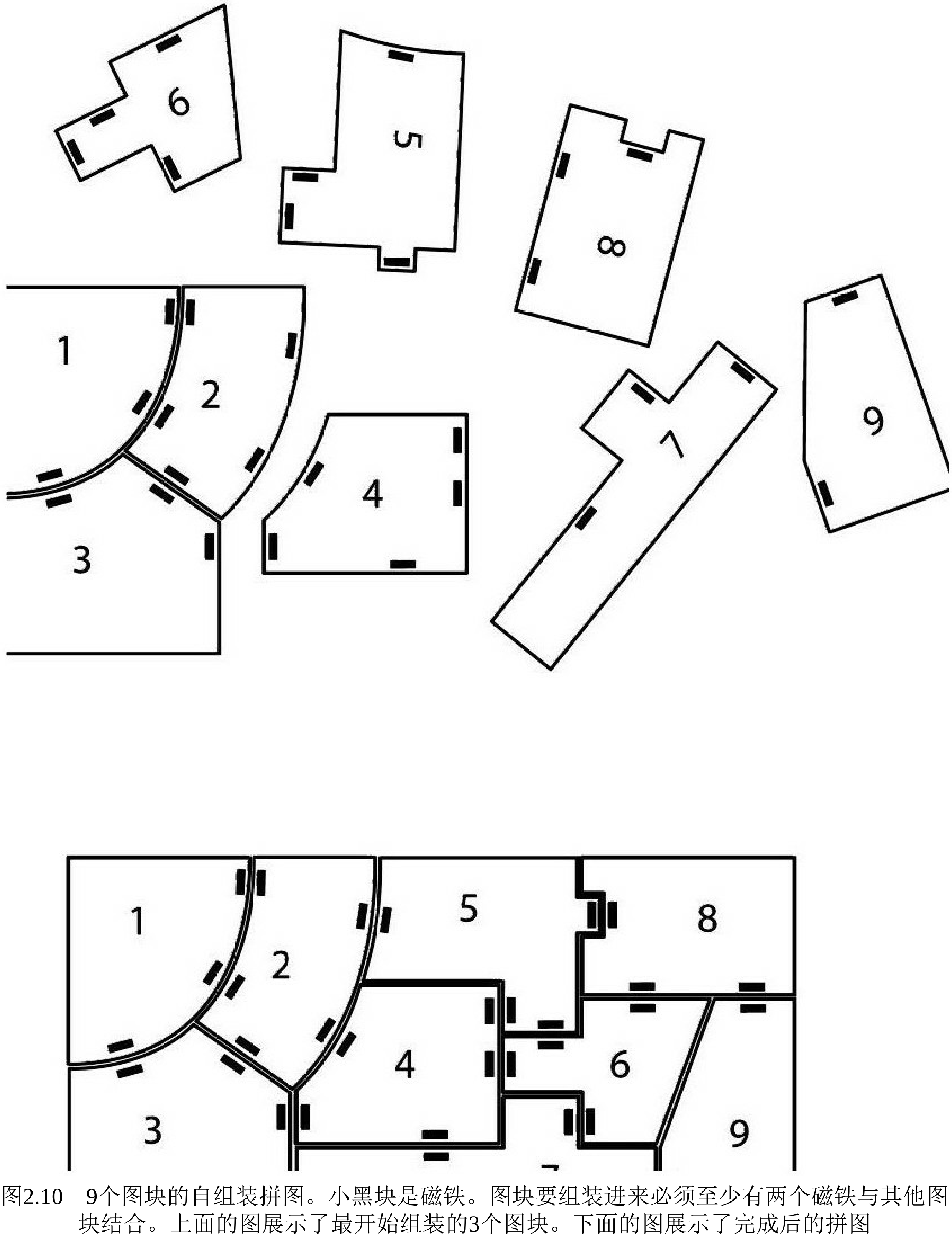


图2.10　9个图块的自组装拼图。小黑块是磁铁。图块要组装进来必须至少有两个磁铁与其他图块结合。上面的图展示了最开始组装的3个图块。下面的图展示了完成后的拼图

在图2.10中黑色的小块是磁铁，摇晃的强度使得单独一对磁铁无法长期绑定在一起；只有当两对或更多磁铁在相同时间接触到一起，图块的结合才会牢固。只有图块1和图块2符合条件，可以无需其他图块的帮助结合到一起。一旦图块1和图块2结合到一起，图块3就能同它们结合，图块3结合后，图块4就能结合，然后是图块5，然后是图块6。图块6结合后，图块7和图块8就都能结合，而图块7和图块8都加进来后图块9才能结合，从而完成拼图。

自组装将通常被区分开的物理过程和计算的概念连接到了一起。最终，物理结构的形成和计算都能通过状态之间随时间的互动来认识。确定性数字计算构成了一个极端，状态只由（0和1）两种组分形成，并且采用的规则总是被遵循（没有错误）。物理系统也是计算，但它们通常有许多组分，并且有噪声，因此错误很常见，即便知道规则也往往无法详细预测确切的结果。

为什么需要长程序？

我们大部分人在思考计算时，不会想到自组装砖块或NKS系统，甚至不会想到任何基于简单规则的物理系统；相反，我们想到的是需要长程序的应用，例如字处理软件和电子表格。根据我们的经验，大部分有用的计算机应用都需要长程序。通过研究一切可构想计算的所有可能的输出，数学家证明了大多数都需要长输入才能产生。考虑到计算和物理组装的密切关联，有理由认为我们想得到的许多事物也会需要长程序。对于NKS和拼砌系统，由于规则和砖块类型有限，这意味着大多数可能的行（输出）如果没有复杂的种子行就无法实现。因此虽然元胞机规则集90、110和1635产生的行很复杂（图2.4、图2.5和图2.6），如果种子行简单，大多数可构想的行永远也无法通过这些计算产生。大部分时候，要得到一个特定的行，你需要从一个仔细设计并且已经很复杂的种子行开始。种子行的细节决定了任何给定规则集的产出。再次强调前面说过的，计算机理论告诉我们，如果输入（规则和种子）简单，大多数可构想的结果都无法通过确定性过程产生。

好的一面是，在符号语言的领域内只要有合适的长程序似乎就能生成一切。经常使用电脑的人对此有经验。长程序能产生令人炫目的输出，字处理、电子表格、网页编辑和计算机游戏只是一小部分例子。大

多数日常计算机应用就是无法用短程序实现，再巧妙的设计也不行。

这个简单的原理几乎肯定对物理世界有效。想一想生命和许多人类技术产品的形成。生物和汽车的产生都需要长而且复杂的指令。根据计算机科学的原理，我们可以想见，如果只依靠化学和物理定律，没有附加的信息，这类事物是根本不可能出现的。这就是指令的作用。

总结一下这一章给出的许多概念，计算可以被认为是逻辑规则作用于已存在的状态产生新状态的过程。至于状态是发生在计算机电路里，还是含有分子砖的试管里，或是活细胞里，抑或是自组装的拼图，并不重要。

系统的状态之间如果可以相互区分，系统就能携带信息。系统状态可以表示其他系统的状态，这一点很重要。这样意义就能从一个系统传递到另一个系统。信息和意义的概念与日常用法纠缠不清。在第1和第2章我们遇到了定义和度量信息的3种方法。用长度度量很简单，但无法体现全部；香农信息是基于状态的概率；算法信息则是基于过程（产生对象的计算）。在确定性计算中长度可以增长，但算法信息和香农信息不会。所有这3种都可以归结为计算的结果。还没有人提出过意义的量化度量。

计算的一个重要原理是显著的复杂性可以通过简单规则反复作用于简单初始条件产生。NKS和DNA拼砌系统提供了实在的例子，化学和物理定律反复作用于物质简单的初始状态形成的物理结构也说明了这一点。下一章我们会看到这种结构的一些例子。

计算机科学的另一个重要原理是一些复杂事物无法通过简单输入和简单规则的计算产生。生命以及许多生命的产物，包括人类社会和技术，似乎都需要长程序。长指令在物理世界扮演的角色就好像长程序在计算世界扮演的角色。没有它们，一些复杂的对象和行为就不可能产生。第4章我们会看到一些这样的例子。

第3章　免费的结构

如果物理过程是计算，程序在哪？

想象一个水分子飘浮在空气中，自由运动，直到碰巧遇到了一片正在生长的雪花。如果条件合适，它可能会粘上去，也有可能会弹开，继续漂浮。这个看似随意的过程不断重复，直到产生出一个美丽而规则的结构。想想流经岩石的河流。水分子无数次地碰撞，形成流向水底的漩涡。再想想规模宏大的氢气云在虚空中通过引力的互相吸引聚拢形成亿万颗恒星和行星组成的旋转星系。最后，想一想上一章讨论过的DNA砖块根据其化学形状组合到一起形成规则的三角形图样。

这些过程有什么共同之处？都是自发形成。都有某种类型的基本成分：原子、分子或分子砖块，根据简单的规则相互作用，形成结构的过程都无需智能的参与。都有某种吸引力法则在起作用。一旦形成，结构之间的互动就会导致新的结构。结构形成结构，成型后又作为种子形成新的结构。如果沿着事件链条回溯，结构之前还有结构，最终模块本身作为规则作用的产物涌现出来。天文学家告诉我们，一切都源自137亿年前。已故的英国宇宙学家弗雷德·霍伊尔在一个广播节目中提出了大爆炸一词，以区别“稳态”理论（他喜欢的）和有明确开端的理论（目前被绝大多数宇宙学家所接受）。根据大爆炸理论，整个宇宙开始于无穷小无穷致密无穷热的点；物理学家称之为奇点。

这就是目前关于宇宙的主流科学观点。自然规律无处不在；它们在所有尺度上，在所有时间和地点都起作用。时间、空间，也许还有规律本身都是起源于大爆炸。规律并没有写在什么地方，但是它们可以被人类观察者发现和描述。

物理定律大部分是数学化的，可以用它们来计算宇宙从最开始又小又热的状态如何演化成现在的样子，但定律并不能解释它们自身的来源。现代宇宙学的圣杯就是找到能解释这些定律的自足的理论。对这个问题感兴趣的读者可以参考注释中给出的文献 [1] 。幸运的是，没有这样的“万有理论”我们也能学习物理定律和研究它们的影响，无论我们是

否能在更深的层面上理解它们，它们都一样有效。

一些定律是确定性的，也就是说，在相同条件下它们总是会导致相同的结果。一些则是随机的，即使条件相同也会导致不同的结果。我们还知道在有些情形下确定性定律的作用会产生混沌，这时无法精确预测长期的结果。由于随机性和混沌，有时候无法准确预测未来的细节；因此虽然遵循的定律是一样的，每一片雪花却都不一样。

第2章我们看到了通过简单计算产生的结构。这些例子表明简单规则作用于简单输入有时候会产生极为复杂的结构。系统的研究表明，许多（可能大多数）简单规则和简单输入的组合都不能产生复杂的结构[2]。许多组合什么都不能产生，一些则产生简单的重复性图样，只有极少数组合会产生出复杂图样。在物理和化学中也是一样，定律作用于简单或平常的初始状态可能什么都不会产生，或是偶尔产生重复性结构，极少产生复杂结构。要了解特定的结构如何产生需要有相关的定律和初始条件的知识。

类似于计算机的输出，物理结构的形成是不确定的，也就是说有许多途径可以形成特定的事物。可能很难发现具体是哪条途径产生所研究的事物。幸运的是，理解一个结构并不需要知道它到底是如何形成的；只要知道它可能的形成途径即可。再次以雪花为例。水分子从不同的角度接近生长的结构，并附着于表面不同的位置；每一个分子的碰撞细节决定了雪花的唯一性，但我们无需知道这些细节也能大体上了解雪花的形成过程。虽然每片雪花的形成细节都不同，但雪花形成的原理基本是一样的。

当考虑形成某种事物的可能场景时，必须承认有随机创造的可能性，组分被偶然地放到一起。结构可以也的确会偶然产生，但随着组分数量的增加，出现特定构造的概率变得越来越小。即使是组分数量不太多的构造，特定的有序结构产生的概率也会小到宇宙的生命期内都不会出现。我称这个为概率问题。如果在某个场合遇到了概率问题，就需要另外的解释。

第1章说过事物可以分成两种结构类型：I型（基于简单规则）和II型（需要复杂规则）。有一种完全不同的对结构分类的方法来自热力学的研究。根据这种分类方法，一些事物的形成是当系统接近热力学或化

学平衡态时，另一些事物的形成则是当系统远离热力学或化学平衡态时。近平衡态结构的例子包括水晶和太阳系。除非有外力改变系统，否则这种结构会一直维持下去。这种结构总是处于特定的平衡态，处于系统的最低能态。远离平衡态结构的例子包括流体中的漩涡，湖中的波浪和太阳。这类结构被认为处于稳态而不是平衡态。例如，生物体就是稳定的非平衡态。非平衡态结构的存在也有赖于力的平衡，但它们的形成只有在有能量稳定流经系统时才有可能。例如当你给某个设备插电或当风刮过湖面形成波浪时。如果能量流停止，就会失去稳定性，结构溃散。简而言之，在平衡态，只有在改变结构时才需要能量，而在非平衡态，维持结构就需要能量。

还有一种分类方法是分为静态结构（不动的）和动态结构（动的）。无论是平衡态还是非平衡态结构都有运动。例如，太阳系中行星围绕太阳转，卫星围绕行星转，就是运动的平衡态结构。这种运动不需要来自太阳或其他地方的持续能量输入。相应的，当你将浴缸中的水放掉时则是[近]静态非平衡态结构，虽然水分子在排下去时在转圈。在水流进下水管时，漩涡的形态基本保持不变。必需的能量流是由水释放的势能提供的。

让事情变得更复杂的是自然界充满了短暂存在的结构。这类事物持续时间不长，形成了系统的暂态。在长期结构和暂态结构之间没有明确的区分。太阳系对我们来说似乎很久远，但相对于银河系的生命却很短暂。在非平衡态结构中，经常有结构从一种形态变成另一种形态，但并不溃散。

让物理结构成为可能的是相反的力的平衡。在特定的条件下物理定律描述的力相互作用。一些物理定律描述物质聚拢的趋势，另一些则描述远离的趋势。在引力和斥力之间自然而然会有平衡点。平衡就会产生结构。我称之为免费的结构 [3] ；意指系统和定律中蕴含了产生结构的机会。无需预先设计；条件满足了结构就自然产生。

下面我们来举两个例子，看看吸引和排斥律的相互作用是如何导致结构的自发形成。两个例子中结构都是通过系统达到热力学平衡而形成，不需要指令。

如何解释一粒盐？

食盐由钠和氯组成。同其他原子一样，钠原子和氯原子也是由组分组成：带正电的质子，带负电的电子，以及电中性的中子。质子和中子在原子的中心被强大的核力紧紧束缚在一起，电子则在外围。我们的宇宙有一条定律是正负电荷互相吸引。引力与带电物体距离的平方成反比，这就是著名的库仑定律。这条定律可以表述为10个符号的简单代数关系式：$F = -k \cdot q_1 \cdot q_2 \div r^2$ q_1 和q_2 各被视为一个符号；点号表示相乘）[4]。这个看上去可能复杂也可能不复杂，取决于你的教育背景，但根据我们之前了解的信息定义，库仑定律并没有表示多大的信息量，它就是一条简单的规则。

库仑定律精确描述了带正电的质子与带负电的电子之间的引力，以及两个质子或两个电子之间的斥力。根据这个定律，大部分原子最喜欢的状态是相同数量的电子和质子尽可能紧密地结合在一起。钠原子有11个质子和11个电子，氯原子则各有17个。如果原子的质子和电子的数量一样多，我们就说原子是电中性的。不过，原子有可能有不同数量的电子和质子。这时原子就具有净电荷，或正或负，我们称之为离子。

根据库仑定律，电荷越接近，作用力就越大。如果只有这条定律起作用，则钠原子中的11个电子和11个质子就会挤碎在一起，也就不会有我们知道的钠原子。这一切之所以没有发生是因为量子力学和被称为泡利不相容原理的规则。量子力学也是可以表示成简单方程的规则 [5] 。量子力学规则的一个特点是系统按“能级”划分。对于原子这意味着电子处于某个能级并且与质子的距离不能低于最低的能级。因为电子具有自旋的特性（可以有两种值），同一个能级上可能有两个电子，各有不同的自旋。规则作用的结果是钠原子的最低能态（平衡态）有两个电子占据能级一，两个占据能级二，两个占据能级三，到能级六就只有一个电子。更高的能级上则没有电子。

量子力学的一个特点是，有2、10、18、36、54和86个电子的原子特别稳定。由于钠的电子数量11，将一个电子拉离原子（变成10个电子）所需的能量很少。一旦失去了第11个电子，就变成了钠离子，具有11个质子，10个电子，和+1的净电荷。氯原子的情形也类似，每个低能级上各有2个电子，能级九上则只有1个电子。由于中性氯原子有17个电子，而18个电子特别稳定，也就意味着氯原子倾向于接受一个额外的电

子。这使得氯原子变成氯离子，总净电荷为-1。量子力学、库仑定律以及数字的相互作用赋予了每种化学元素独有的特征。

当等量的钠原子和氯原子混合到一起，就会发生自发的反应，使得系统转变成新的平衡态。这个化学反应会释放热（能量），每个钠原子失去一个电子，每个氯原子获得一个电子。根据库仑定律，带正电的钠离子和带负电的氯离子相互吸引，但不能距离太近，因为填充的能级（化学的轨道）不能相互渗透。当所有钠离子都被氯离子包围，氯离子也被钠离子包围，形成非常规则的3维结构，相抵触的规则之间就会达到最优平衡。能级的物理维度决定了离子之间的最优距离。钠离子和氯离子具有不同的大小，能平衡物理和几何不同需求的状态是一个规则的网格，每个氯离子周围有6个钠离子，每个钠离子周围也有6个氯离子。这个规则的立方体结构向各个方向延伸，从而形成我们所熟悉的盐晶；在适当的温度和浓度条件下钠离子和氯离子相遇就会自发形成这种结构。当钠离子和氯离子溶解在水里并且水缓慢蒸发时这尤其可能发生。

盐晶是室温下钠和氯离子的低能（平衡）态。维持它不需要能量流，结晶过程中也不需要外来的能量源。晶体是免费结构的简单例子。3个简单规则——库仑定律、泡利不相容原理和量子力学——的相互作用决定了其结果。盐晶的形成可以通过有合理可能性的初始条件（例如，溶解于水的钠离子和氯离子以及逐渐失去水）和3个规则得到解释。其他化学成分的存在可能让情况变复杂并导致其他结构的形成；但适合盐晶形成的条件并不是那么罕见，岩盐（氯化钠晶体）在地球上也是很常见的矿。

结晶是计算吗？在第2章我们看到，DNA砖块根据简单规则的结合就能“计算”结构。在这两种情形中都没有常规意义上的算法，但都是规则作用于初始结构（输入）产生出了最终的结构（输出）。这与第2章图2.4中用元胞自动机的计算创造出各种大小的三角形从而形成结构有多大区别吗？在这些例子中都不存在明显的程序，作用于结构的规则决定了发生的事情。

太阳系是怎样形成的？

运动的平衡态结构的一个例子是太阳系。我们的太阳系大部分也可

以视为三个作用定律的产物，分别是能量守恒、动量守恒和引力。能量守恒定律说的是能量既不能创生也不能消灭，只能转换形式。常见的例子包括化学能转换成火的热能，跳伞时势能转换成动能，发电机将机械能转换成电能。动量守恒说的是运动物体除非受到外力作用，否则会一直保持运动速度和方向。引力定律的经典形式说的是两个物体会互相吸引，吸引力正比于两者质量的乘积，反比于两者距离的平方。引力定律的表达式类似于库仑定律，只不过吸引的是质量而不是电荷，并且质量从不互相排斥[6]。

如果不在意细节的话，太阳系的形成很容易理解。故事大致是这样。在宇宙中不断有恒星诞生。在太阳出现之前，空间中有气体和尘埃组成的云。云中大部分是氢和氦，但也包含碳、氧、氮、铁等在现在的太阳系中找得到的元素。氢和氦大部分在宇宙大爆炸最初的几分钟里形成，其他元素则在恒星中形成，恒星在超新星爆发后将其组成物质抛散到空中变成尘埃微粒。这些超新星很重要，因为它们创造了重元素，并且爆炸产生的物质流会导致附近的尘埃和气体云运动，分布变得不均匀。

空间中不均匀的气体云是不稳定的。引力定律决定了所有原子、分子和尘埃会相互吸引。结果原子向原子运动，物质块向物质块运动。动量就是对这种物体相对运动的度量。

对于运动的粒子，如果最初没有朝质心运动，动量守恒会阻止它们直接相向运动。动量守恒和引力定律的交互作用会使得粒子围绕高密度的中心旋转。

能量也必须守恒。当粒子相互旋转时，它们的速度会增加，因为引力势能会转换成动能。随着相互吸引的气体和尘埃粒子旋转得越来越近，它们的速度也会越来越快，发生碰撞越来越频繁。碰撞后会发生什么取决于粒子的化学性质。一些碰撞是弹性的，粒子就好像台球相互弹开。一些碰撞是无弹性的，粒子就好像橡皮泥一样粘到一起。无论哪种碰撞都必须遵守能量守恒和动量守恒。

碰撞还会产生热，以红外辐射的形式放射。粒子碰撞、弹开和黏合到一起的结果是黏合的粒子越来越多，并且所有粒子的运动变得协同一致。所有粒子聚拢并降低速度，导致热能损失。由于旋转，气体云会变

成盘状。大部分尘埃和气体会在中心附近，少量在外围。能量和动量守恒使得中心附近的粒子和气体运动得最快，最外面的物质则运动得最慢。随着分子越来越接近旋转中心，碰撞也越来越频繁。碰撞越多释放的热量也越多，中心也变得更加热和致密。最终云团中心变得如此热和致密，以至于开始出现热核反应。这又会产生更多热和辐射。辐射将原子和分子外推，与向内的引力相互抵消。

当辐射导致的外推与引力的向内拉相互平衡，就会形成稳定的结构；我们称之为恒星或太阳。这时外部的气体和尘埃盘变得引力不稳定，因为大块环绕物质的引力会改变附近物体的轨道。改变的轨道相交，导致的碰撞形成更大的物体，对附近的聚集体、气体和尘埃的引力也越来越大。变大的物体相互吞噬，最后只留下少数大的聚合体（行星）。行星围绕太阳形成稀疏有序的结构，相互之间的距离遥远，不会相互干扰。如果它们靠得太近，就会相互拉扯偏离轨道并最终相撞。

大致就是这样。实际的细节极为复杂，但三个简单规则的交互确保了，只要存在规模合适的、运动的和不均匀的气体尘埃云，就会形成有序的太阳系结构。太阳系的具体细节取决于初始气体云的大小和成分"黏合"规则（量子力学）和热核反应的规律。不过只要考虑三条简单规律——引力、动量守恒和能量守恒——就能大致理解太阳系有序结构的形成。由于宇宙中有大量的气体和尘埃云，可想而知太阳系必定很常见，虽然最近才开始对它们有所认识。运动在太阳系中扮演了重要角色，因为只要物体在运动，就要遵循动量守恒。只要质量块不是直接相向运动，动量就会对抗引力，行星轨道就是沿直线运动的趋势（动量）和相向运动的趋势（引力）之间平衡的结果。

太阳系体现了两种结构。行星的稳定轨道是平衡态结构。行星的持续运动和分布不需要能量输入，只要没有摩擦和外力，太阳系会永远运转下去。太阳中心的结构则不同。如果核反应堆耗尽，太阳就会缩减到目前直径的百分之一。如果质量足够大，就会变成黑洞。恒星需要热核反应提供的持续能量来维持其稳定结构。它是非平衡态结构。

非平衡态结构如何形成？

不均匀加热的流体和气体是另一种自发形成结构的例子。这种结构

比吸引和排斥的简单平衡形成的结构要难理解一些，但也更为引人入胜。

烧水就能观察到非平衡结构。如果从底部加热一盆水，就会发生惊人的变化。如果加热不剧烈，不会发生什么值得注意的事情。在亚微观层面上，与盆底碰撞的水分子会加速反弹，通过与活跃振动的盆底原子碰撞获得动能。能量通过碰撞在水中传递。如果加热不剧烈，表面的水会比底部的水温度低，但不会低很多。如果提高供热，更多热量会被传递给盆底，剧烈的碰撞会增多。同时表面和底部的水温差别会增大。

热水比冷水的密度低，因此当加热到一定程度时，就会出现高密度的冷水在上面，低密度的热水在下面的不稳定情形。当容量相同，高密度冷水会比低密度热水受到更大的地球引力，如果受力差别大到一定程度，上面的冷水就会通过对流和热水交换位置。水分子在液体中不停的碰撞，因此活动并不那么自由，而是协同运动。在一些区域，水会一致往下流，另一些区域则会一致往上走。在特定条件下，从底部往表面传送热量的最佳方式是一种被称为对流环的协同运动。如果液体很薄，就可以观察到特别有趣的情况。这时表面张力压制了浮力，如果盆很平，热量传递很均匀，就会自发形成非常规则的六边形对流环，这被称为贝纳德流（图3.1）。在这种结构中，水从每个环胞的中心上升，在环胞边缘下降。如果条件没有严格满足，环胞就没那么规则。如果将液体限制在两块平板中间，顶部不开放，就会形成翻滚（称为瑞利—贝纳德卷）而不是六边形环胞。

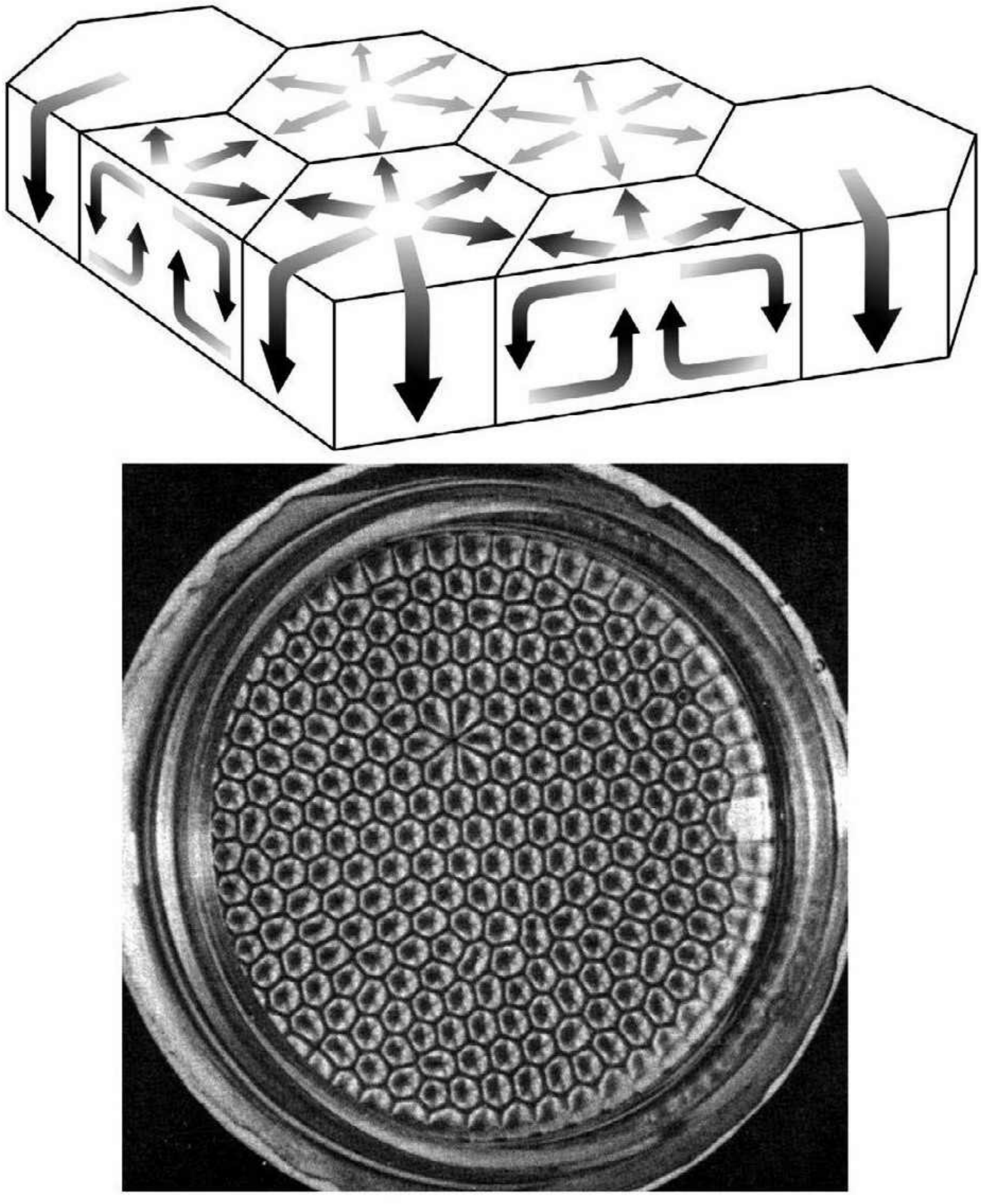


图3.1　贝纳德流的图和照片。图中给出了几个六边形对流环胞的截面图。箭头为运动方向。照片是从顶部观察薄层硅油，为了便于观察添加了铝粉。照片经科施米德尔许可引用

这些规则运动很稳定，只要对盆底的供热量合适（不太多也不太少）就会一直保持。供热太快，就会变混乱。供热太慢，协同运动就会

停止。

对流环胞是非平衡态结构的经典例子，大量分子自发协同运动，无需指令告诉它们怎么做。这个特性使得一些科学家提出远离平衡态的自组织有可能为生命系统的许多复杂性提供解释。这个问题还悬而未决。贝纳德流还不是特别复杂。同晶体类似，它们也是源自简单规则的交互作用。对于上表面敞开的薄液层（贝纳德流所需的条件），起作用的是黏性、表面张力（分子之间的吸引力）和热运动（扩散）[7]。对于较深的液体，浮力（基于密度和重力）压制了表面张力驱动对流。

有一个基于对流的更复杂的结构是雷暴。暴风雨在许多方面类似于贝纳德流，不过是发生在大气中而不是液体中，浮力是驱动力，只是环境不像火上的平底盆那样规整。大气对流还有其他特点，空气上升会变冷，密度降低，当温度降到凝点以下水蒸气会凝结成水。同时水凝结时会释放热量。上升的湿空气凝结产生的热量会让空气升得更高，从而进一步冷却。

热带风暴开始于晴热的早上，低空大气中有许多水蒸气。太阳光加热地面或海面，反过来加热低空大气。低空大气逐渐变得越来越热。最终大气变得不稳定，底层的热空气上升，上层的冷空气下沉，很类似贝纳德流，但水蒸气的存在让事情变得不同。根据理想气体定律，上升的空气扩张冷却。当冷却到一定程度，水蒸气凝结成水滴，释放热量。释放的热量加热空气，导致其进一步上升，并继续扩张冷却；这反过来又导致进一步冷凝和释放更多热量。空气上升后，附近的低层温暖湿空气流入补充，产生更多水滴。小水滴碰撞聚集成大水滴。

最终，上升的空气无法支撑不断增大的水滴，水滴开始下降。水滴在下降过程中会裹挟空气分子产生向下的气流和雨滴。水滴会聚集自由电子，带有水滴的风在云的不同区域产生大的电离，放大了云层和地面的电位差。当电离大到一定程度，就会通过集中的电流中和，也就是我们看到的闪电。这种电流会让空气局部剧烈升温，引发声光现象。含有水的空气的上升下降也会产生水平方向甚至旋转的风。风暴通过不断补充湿热空气和降雨（冷到一定程度就降雪）排出冷水维持。一旦没有了上升的湿空气，风暴就结束了。这一般是因为太阳下山了，也有可能是因为流入的是干空气，或者风暴遇到了高空暖气流。

类似于晶体，贝纳德流和雷暴也是少量简单规则作用于不是特别罕见的初始条件形成的。

结构必须有永久部分吗？

风刮过水面会形成波浪，风越大，浪就越高。水波是非平衡态结构，它们的形成有些类似于贝纳德流，不过能量的来源是风。类似于雷暴，形成波浪的分子也只是暂时参与。它们短暂地成为一个长期结构的一部分，结构离开时并不带走分子。从这个角度来说，波浪和雷暴独立于它们的组成成分（分子）。许多结构都有这样的特点，包括我们自己的身体。为何会这样？

水波提供了一个很好的例子。当风刮过平坦的水面，空气分子与水面的分子碰撞，导致它们运动。液体分子相互牵制，因此无法运动很远。由于微观的不均匀性，风的分子与部分水分子的接触更多。这导致它们“堆叠”，就像贝纳德流一样，这种运动是大量分子的协同运动。最初的堆叠只是涟漪，一旦形成涟漪，迎风面的分子受到的风力就会大于背风面。由于分子不断相互碰撞，能量沿着水面在分子之间传递。由于水很难被压缩，每次运动都必然伴随着补充运动。在各种限制下，自发出现的协同运动是分子沿垂直方向的环运动。大量垂直的环形运动的总和就是波浪。随着波浪经过，分子前下后上转圈（图3.2）。

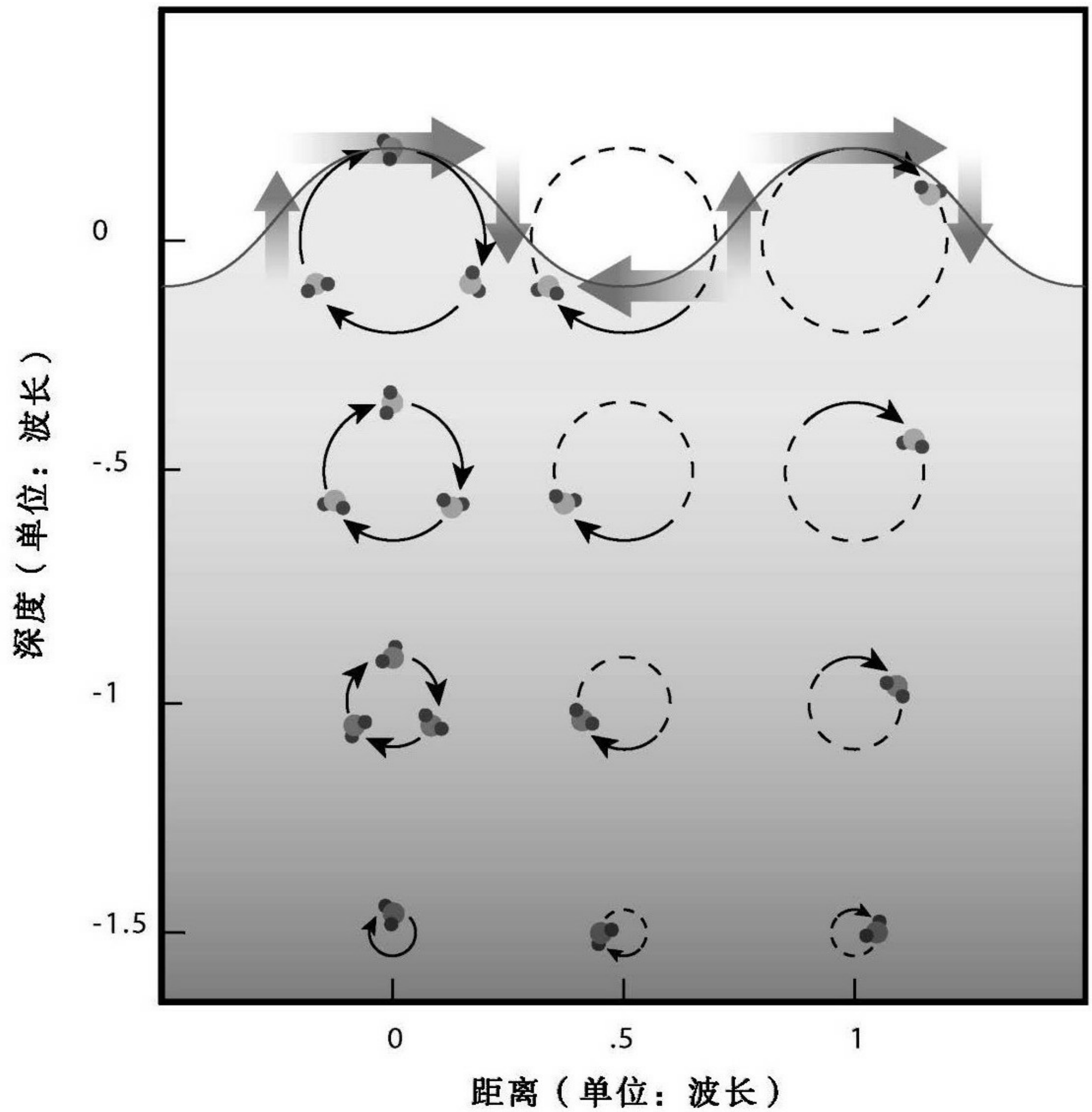


图3.2　波浪中的分子协同运动。箭头标示了波浪不同部分的运动方向。波浪从左往右运动，水分子则沿环形运动

位于波浪顶部的顺着风向运动，位于浪中间的则反向运动。所有分子都有节奏地前下后上运动，但不会运动很长距离，虽然波浪可能会走得很远。这种协同满足了随风力运动和“补充”浪顶向前运动的水的要求。表面的分子没有离开表面；水面则随着分子的环形运动起伏。表面下层的分子也作环形运动，但越深圆环越小。

大量水分子高度协同的运动组成了动态非平衡结构。风刮得越大越

久，环就越大，产生的波浪结构也越大。一旦风停了，有组织结构就会逐渐耗散，波浪也越来越小。运动的计算很复杂，但决定波浪形成和推进的规律很简单，只不过是黏滞性（水聚在一起）、不可压缩性、风传递给水的能量以及水分子之间的能量传递（守恒）。分子对波浪结构的参与是暂时的；它们参与一次循环然后退出。结构本身就好像是一个对象。这里同样是简单规则作用于平常的初始条件就自然形成结构。

行为是一种结构形式吗？

物理事物由规则作用于3维空间的物质产生；行为则由规则作用于时间的结构产生。水波有节奏的变化，方式很简单。如果系统可以储存势能并间歇释放，则能产生更复杂的行为。一些会有不规则但有特点的行为。闪电就是这样的例子，另外还有陡峭山坡上不稳定的积雪。一旦大意的滑雪者震动了雪堆，积雪就会通过一系列复杂的运动沿山坡崩塌。储存的势能通过这种运动释放，然后再次累积至临界态。

物理学家在实验室用沙子研究了产生崩塌的系统的行为。结果发现，仅仅只是缓慢堆积沙子，其他什么也不做，沙堆也会表现出复杂得惊人的行为。随着沙子逐渐累积，沙堆越来越高，坡度也越来越陡。最终，再增加一点沙粒就会导致崩塌。继续堆积沙子，会重建沙堆，直到再次发生崩塌。有趣的是，崩塌的规模大小不定。有时候会出现大规模崩塌，但规模越小的崩塌越频繁。如果仔细测量，会发现崩塌的规模遵循所谓的“幂律”[8]。发生大规模崩塌的可能性要比小规模的小得多。根据幂律，事件规模每增加一倍，特定规模的事件数量会减少一个常量。图3.3给出了一个例子。幂律关系表明，崩塌的规模看似随机，实际上却以一种有序的方式相关。这种相关性可以视为一种结构。

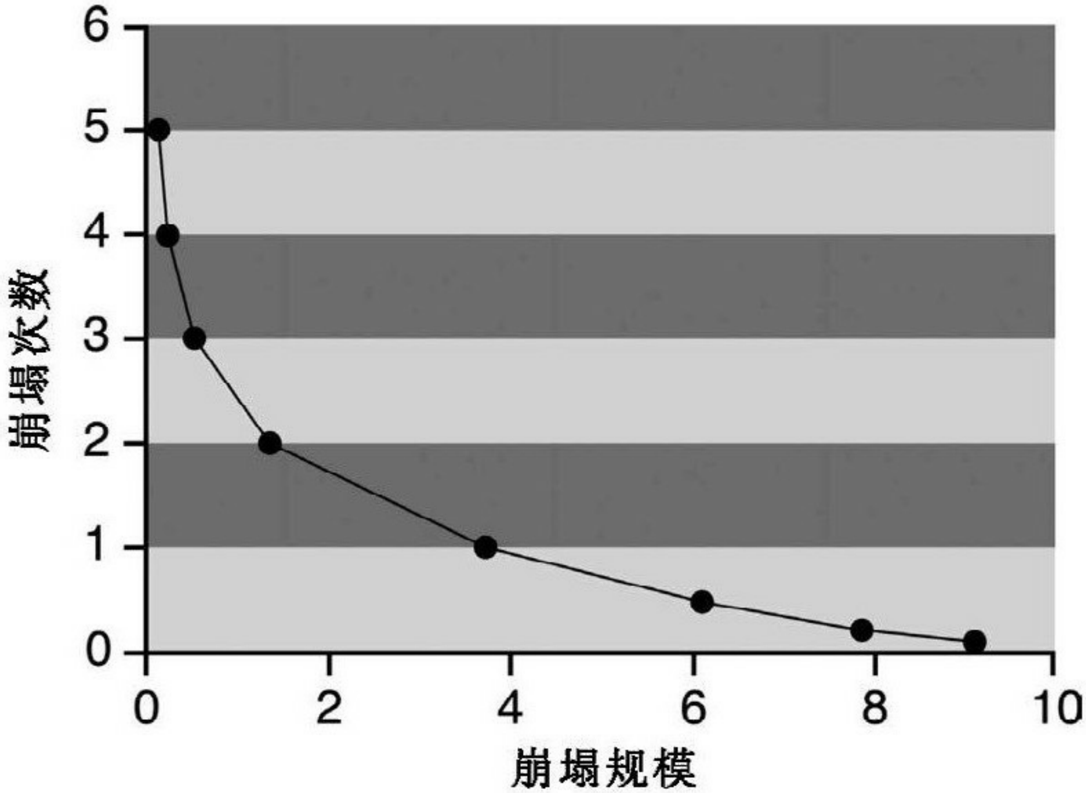


图3.3　崩塌次数和崩塌规模之间的非线性关系，小规模崩塌比大规模崩塌要频繁得多。如果用对数坐标绘制，得到的图会是往右倾斜的一条直线

许多现象都表现出幂律。其中包括许多看似不相关的事物，例如地震的大小和频度，挪威海岸线上峡湾的大小和数量，英语词汇的使用频度和排名，生物灭绝的规模和频度，以及城镇和城市的大小和数量。佩尔·巴克在《大自然的运作原理》一书中对此有更详细的讨论[9]。巴克观察到表现出幂律性的系统具有一种特性，他称之为自组织临界性（SOC）。这个词指的是在有序（比如晶体）和随机（比如气体）之间保持平衡的一种状态。在这些特殊的边缘区域，化学和物理定律会导致特别复杂的状态，系统中的所有元素都相互影响。这种特别的状态的发生比人们预期的似乎更为频繁，因为它们是“自组织的”。也就是说有些系统会随着时间自然发展直至达到一种临界状态。沙堆提供了一个很明显的例子；临界态就是再堆一粒沙就有可能导致崩塌的平衡点。这种系统不能越过临界态，因为崩塌会让其处于亚临界；再堆一些沙子又会回到临界。具有这种特性的系统有可能会表现出幂律行为，从而以一种可预测的（也就是结构化的）方式行事。佩尔·巴克等人认为自组织临界性在许多复杂行为中都扮演了重要角色。

毫无疑问自组织临界态是一些非平衡态系统的特点，但目前仍不清楚自组织临界态对于复杂性的全面认识有多重要。就好像开车，你的里程表从1开始增加，直到10，里程表的第一位变成0，第二位增加1。这是一种崩塌，第一位记录的前9千米被舍弃，代之以十位的1。这个模式不断重复，直到100千米，这时会发生更大的崩塌，前两位的记录被舍弃，代之以第三位的1。如果将里程表翻转的规模与各种规模的翻转次数的对数描出来，就会得到一条表现出幂律行为的直线。里程表表现出某种自组织临界性，但并不能给复杂性问题带来多少认识。沙堆的例子吸引人的地方在于，简单系统通过表现出不同规模的崩塌能产生显著的行为复杂性。崩塌的发生在统计上具有可预测性，但除非你跟踪沙堆里的每一粒沙子，否则具体的一次崩塌是不可预测的。

崩塌可以用计算机模拟吗？

在第2章我们看到，在计算和物理过程之间没有明确的界线。同样，沙堆的思想也很容易通过计算机来体现。同物理沙堆一样，算法沙堆产生的复杂结果也具有统计幂律性。

图3.4给出了一个例子。原理如下。计算机程序在一个有25个（虚拟）格子的“板子”上随机堆积（虚拟）沙粒，图中每一个大方块就是一块板子，填有数字的小方块就是格子。每个格子能堆3粒。一旦堆到4粒，就会散落出去，周围4个格子每格1粒，原来的格子变成空的（图中的0）。如果板子边缘的格子有4粒，1粒就会从板子边缘“掉下去”，离开系统。

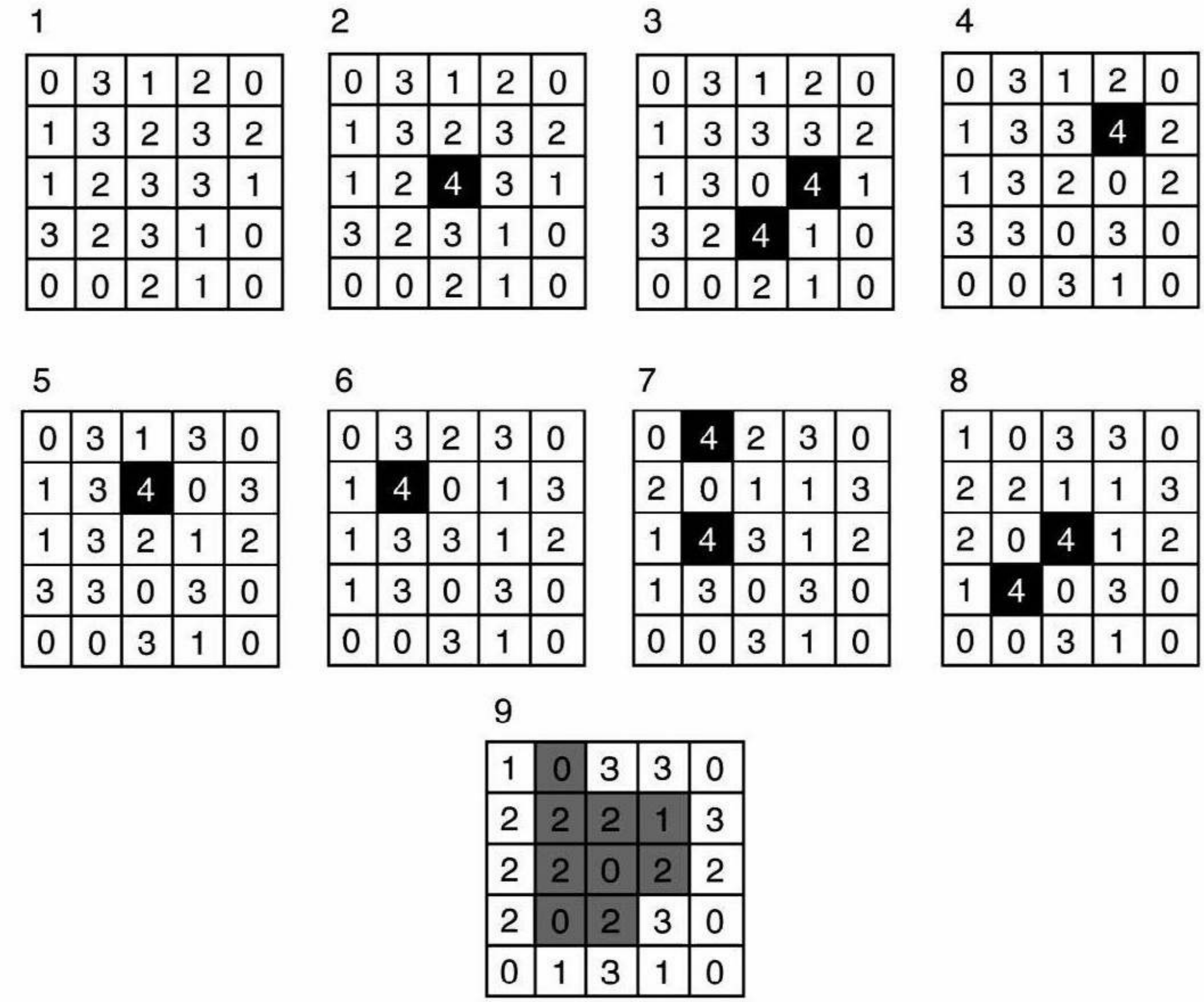


图3.4　崩塌模型。在板子2中间的格子堆了第4粒沙。然后4粒沙散开到相邻的4个格子。这导致又有两个格子有了4粒沙（板子3）。这些沙再次散开到相邻格子。崩塌一直持续到没有格子有4粒沙。板子9给出了这次崩塌最后的状态和所有受影响的格子（阴影）

从空板子开始（全0），沙粒逐渐累积，直到一些格子有3粒，如图左上的板子所示。如果3粒沙的格子再堆1粒，就会触发一系列动态过程，具体的发展取决于板子上所有格子的状态。在图3.4的板子2中，中间格子堆了第4粒沙。4粒沙散开，导致两个相邻格子有了4粒沙（板子3），继续散开又导致另一个格子有了4粒沙（板子4）。这个过程一直持续，直到所有格子都没有4粒沙（板子9）。这个过程就是崩塌，这次崩塌包括9个格子的10次散落事件。这个过程中，中间格子散落了两次。板子9中没有格子有4粒沙，并且标出了所有曾发生散落事件的格子。

堆一粒沙导致的后果取决于堆放的位置和板子上所有格子的状态。如果沙粒被堆放到没有3粒沙的格子，则不会发生崩塌；如果沙粒堆放的格子有3粒沙，但相邻格子都没有3粒沙，则崩塌规模为1；如果相邻格子有3粒沙，则会发生更大规模的崩塌。如果生成随机初始状态的板子，反复玩这个游戏，并描出崩塌规模与崩塌次数的对数关系，结果将是一条表现出幂律的直线。崩塌规模的上限由板子大小决定。板子越大，能产生的崩塌就越大，但总是越大的崩塌发生的频度越低。

这个基于计算机的系统再次阐释了什么是自组织临界性。除非有格子堆了3粒沙，否则就不会产生崩塌。所有格子都不到3粒沙的板子是亚临界的。一旦有格子有了3粒沙，系统就是临界的，就有可能发生崩塌，但崩塌何时发生，规模多大，取决于系统的具体状态和下一粒沙落在哪里。系统不能越过临界态，因为4粒沙的格子是不被允许的。因此，只要有新的沙粒落到板子上，系统就必然“自组织”成临界态。结果的幂律分布反映了导致特定规模崩塌的状态的潜在可能性。也就是说，可能产生大规模崩塌的状态要比可能产生小规模崩塌的状态出现的可能性低得多。

这个系统证明了，幂律型复杂性可以通过简单规则作用于特定初始状态产生。计算沙堆和真实沙堆唯一的区别在于“粘连”规则不一样。对于真实沙堆，是物理定律决定了沙粒在掉下来前可以堆多高，而对于算法沙堆，则是程序决定了一个格子可以堆多少粒沙。

免费的结构从何而来？

结构和行为可以通过简单规则作用于特定初始条件自发形成，第2章讨论的《新科学》和拼砌系统，以及这一章讨论的盐晶、太阳系、贝纳德流、雷暴、波浪和崩塌，只是其中一小部分例子。不是所有条件都会产生结构，但有一些会。无论是计算机中的程序，还是平衡态物理结构，或是非平衡结构，原理是一样的。这些细节不会为我们从总体上理解这些结构如何形成带来多大区别。它们都揭示了一个结构形成的基本原理：一旦适当的简单规则作用于特定的情境，结构就会作为交互作用的规则的结果而产生。这并不意味着只要是简单规则和情境的结合就能形成结构——大部分都不会——但能产生某种结构的规则组合也不是那么罕见。它们存在于自然界，并且通过人工系统得到也不是很困难。

对于这一章讨论的6个系统，如果从微观层面上去跟踪结构形成过程中单个原子或分子的行为，细节会非常复杂，但如果后退一步，从宏观上来研究其行为，就会看到只需少数规则交互作用于特定但不罕见的初始条件就会导致结构的自发形成。贝纳德流、波浪和雷暴都需要通过系统的持续能量流；晶体和太阳系则不需要。对于雷暴，是附近大气的异常导致了结构的不断变化。当沙粒堆积到沙堆上，就会产生统计上可预测的结果，虽然具体的时间很难预测。沙堆算法表明不可预测性来自沙粒之间的大量互动。在临界态，系统所有部分都相互影响。一旦出现这种情况，行为就变得复杂。

这一章讨论的所有例子的特点是：其作用的简单规则和条件都是有可能的。值得注意的是，尤其是对于非平衡态结构，在很短的距离内相互作用的分子（10^{-8} cm）形成的结构和行为的尺度从纳米级到数亿千米都有，取决于系统；并且都无需任何的预先计划。没有蓝图；只需简单的初始条件以及化学和物理定律就可以。这是真正免费的结构。

在这一章的开头我提出了以下问题：如果物理过程是计算，程序在哪？反思这些例子，都是物理定律作用于初始条件产生新的结构。因此，初始条件以及物理和化学定律就是“计算”的输入。不存在预先设计的程序，也没有指令，只有作用的规则，但这些输入就足以产生显著的结构。

下一章我们将继续讨论依赖于物理定律的结构，但不同的是它们需要附加的，并且经常是很复杂的局部规则。这些规则的概率极低，不可能通过随机的方式产生。它们需要特别的解释。

第4章　目的性结构

用不免费的指令能得到什么？

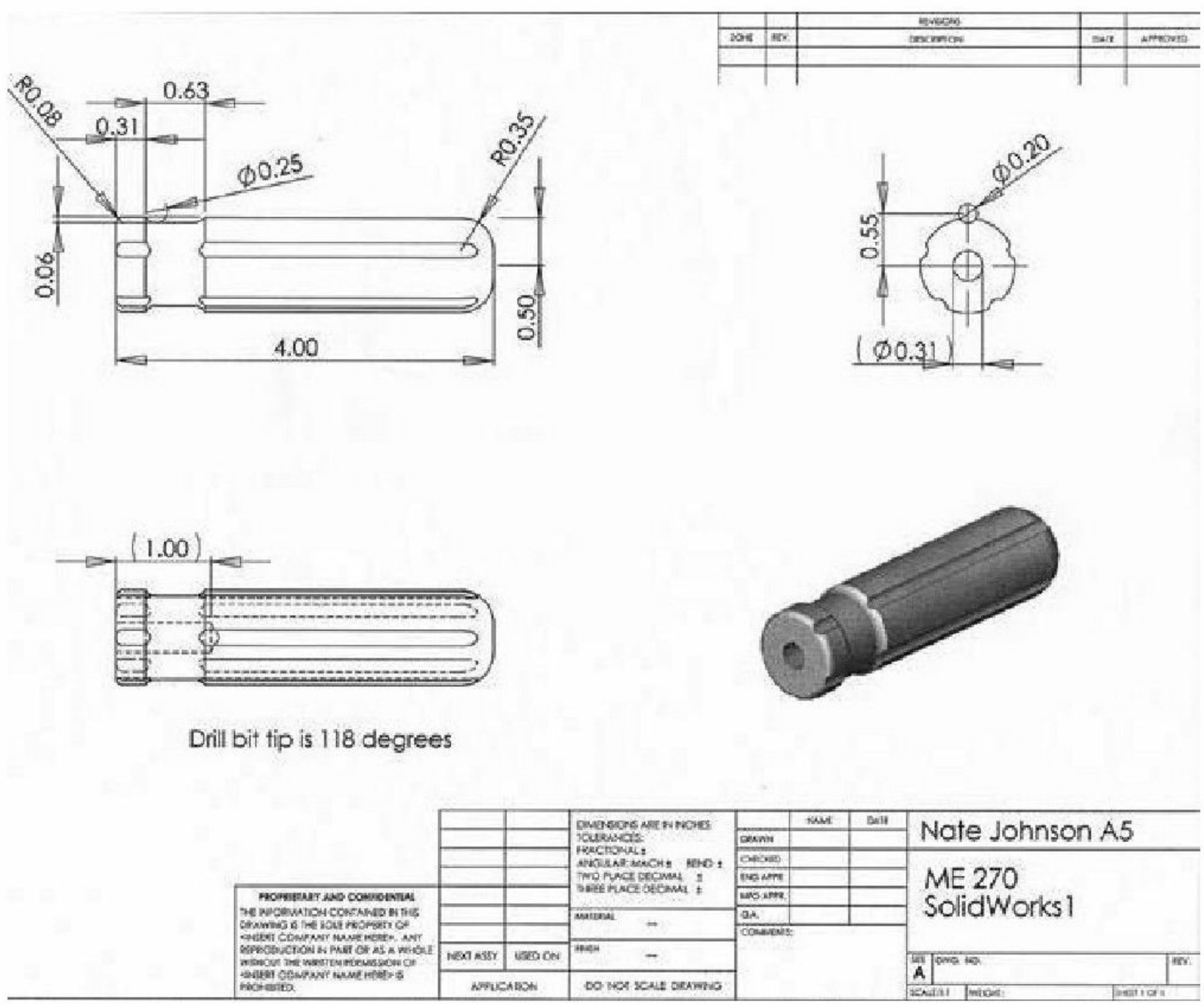
假如你是一位发明家和企业家，想到了一个创意。你用木构件和绳子制作了一个原型，看上去似乎能行。你该怎样做才能用铝合金和塑料大量生产呢？你不会去请一位理论计算机专家替你设计只需摇晃就能自行组装成产品的简单构件。你也不会去请一位纳米工程师替你设计微型机器人，只需简单编程就能放到装有铝和塑料块的箱子里，自行将构件组装成产品。你要做的是画出设计图，写出熟练机工和模具工能明白的指令，以及如何组装零件的指令。有了合适的指令，你就能制造这个新产品的许多份拷贝。

上一章我们看到简单规则（化学和物理定律）作用于简单条件就能自发形成结构。我称之为“免费的结构。”包括雪花和太阳系在内许多神奇的事物都是这样形成的。但是免费结构的概念（短程序反复作用于简单输入）无法解释草木、贝多芬交响曲和计算机。

世界上有许多事物无法用简单规则作用于简单初始条件产生。从计算的角度来说是需要长程序。从物理的角度来说，长程序的形式可以是用于完成特定任务的指令、方法和设计图。它们编码目的性信息。指令并不限于人类活动，可以是任何指导事物形成的规则。只要遵循这些规则，就能得到预先设定的结果。结果具有某种目的性，并且生成结果所需的特定规则（指令）不同于一般的物理定律，这些规则只在某种背景下起作用。例如，DNA（脱氧核糖核酸）和RNA（核糖核酸）编码的指令只在活细胞中运作，设计图只对机工或建筑工程师有意义。指令并不是要替代化学和物理定律，而是利用它们，并且它们总是具有目的。我们以一种常用工具的制造过程为例，来看一看这种信息的特殊之处。

很少有工具像螺丝刀一样简单而又实用，做事的人都少不了它，但很少有人会自己做一把。虽然很常见，但所有人都知道螺丝刀是人类活动的产物。我们可以肯定，没有哪把螺丝刀是自然定律自发作用于天然形成的组件的产物。为什么呢？让我们看看制造螺丝刀的细节。

大二工科学生内特·约翰逊在艾奥瓦州立大学的机械工程入门实验课上做了这种工具。螺丝刀由两部分组成：一根硬杆，一端是平的，以便插入螺丝头部的槽，以及直径粗一些的手柄，适合握在手里。粗手柄利用杠杆原理，这样用适度的力转动手柄就能在平头部分产生很大的扭矩。刀头有多种形状，但这不会改变螺丝刀的基本特性。图4.1是内特在课程作业中的设计图和计算机生成的效果图[1]。内特和他的同学拿到的原材料是8毫米粗的钢杆和25毫米粗的塑料棒。还有其他一些材料。杆和刀口必须坚硬，但不能易脆。钢最适合，铜或钛金属也可以。手柄可以用木头或陶瓷。在课程中学生会学习使用复杂的电动工具，但他们也可以用手钻、手锯、锉刀、刻刀（用于手柄造型），甚至可以用石头当锤子。



杆和到头

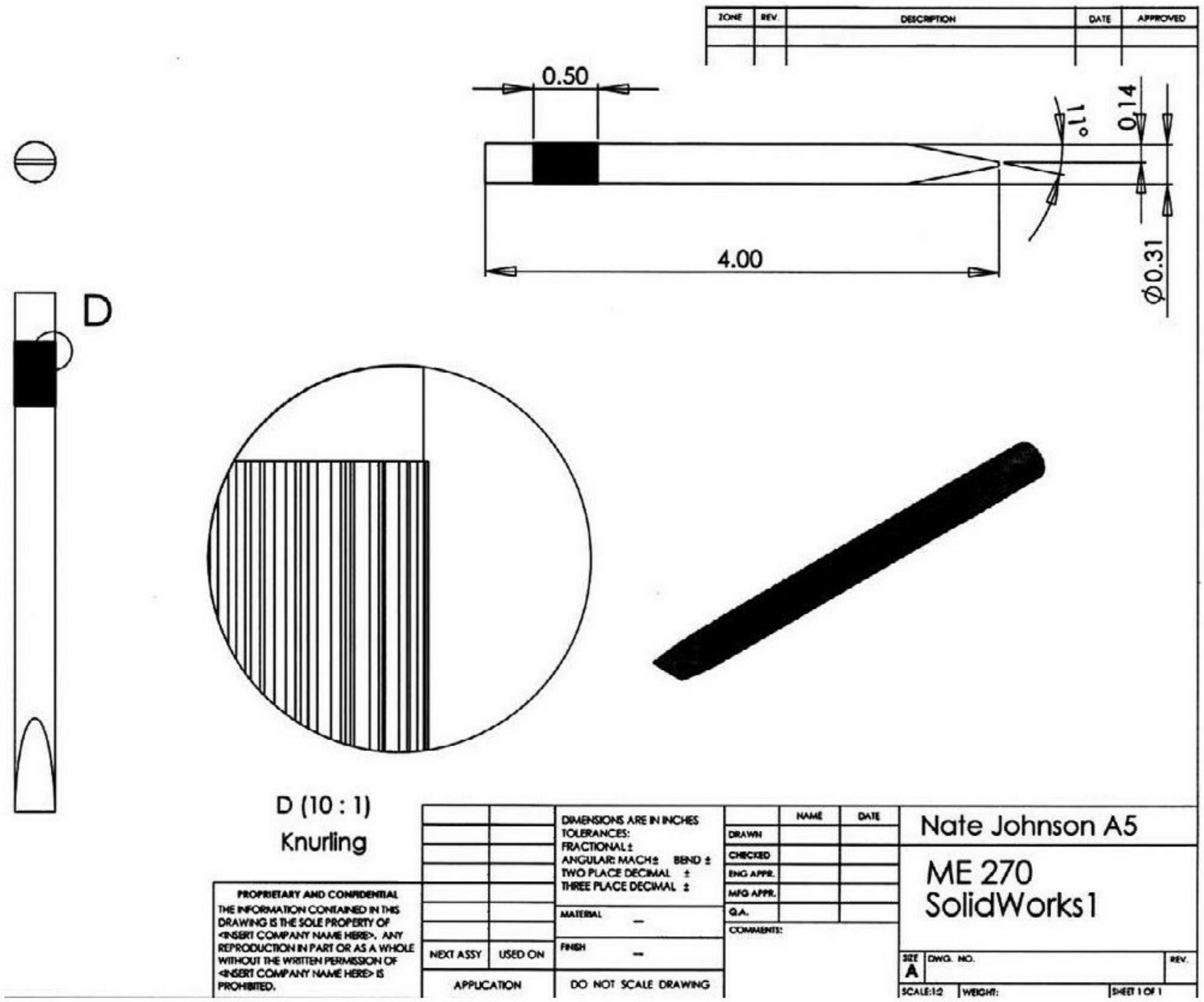


图4.1　螺丝刀设计图和计算机生成的效果图。经内特·约翰逊许可引用

设计图为制造者提供视觉指导，但要利用设计图内特还需要设计一系列制造步骤。表4.1列出了制造步骤，遵循这些指令，就能制造出图4.1中描绘的螺丝刀。

在制造螺丝刀的过程中，同盐晶的形成过程一样，内特也依赖化学和物理定律。两者都是从初始状态开始形成结构，但与盐晶不同的是，螺丝刀的结构不会自发形成，概率太低了。指令的作用是说明一系列低概率的状态。在这里指令表现为步骤，每一步都设置一个状态，生成的结果作为下一步的条件。设计图的阅读者必须将图转化成步骤。而内特作为制造者也是这些状态的组成部分。一旦设定，状态就会引发一系列完全由化学和物理定律决定的事件过程（免费的结构）。

我们仔细来看第一步。钢杆本身就是很特殊的状态，在离一端10厘米的地方被压在运动的锯片上。离端部的距离与物理无关，但一旦钢杆被压在迅速运动的锯齿上，结果就由物理定律决定。锯齿在钢杆表面撕扯。单个金属原子之间的化学键被机械力拉开，将金属屑从深切槽的表面切削下来。具体的细节取决于施加的力、金属的硬度以及锯齿的硬度、锐度和速度，但只要金属棒被压在快速运动的硬金属锯齿上，就必然会挖出锯齿宽的槽并最终截断钢杆。一旦杆被压在运动锯齿的切削面上，初始状态被设置好，之后的变化就取决于化学和物理定律。这部分过程是免费的结构。

表4.1中大部分步骤都涉及各种切削工具的使用，只是初始配置状态不同。最后一步有点不同。内特将钢杆的非刀口端插入手柄上钻出的8毫米粗的孔，并捶打手柄的另一端。动量和能量守恒确保力从锤子传递到手柄，导致钢杆改变孔的内表面。如果力足够大，变形会导致塑料挤在钢杆被压扁处，手柄就会和钢杆固定在一起，螺丝刀就做好了。锤子施加的力度很关键。力太大钢杆会击穿手柄；力太小钢杆又和手柄结合不紧。制造者可以通过反复敲打来调整总共施加的力。重要的是一旦钢杆被放入手柄的孔中，并且挥动了锤子，之后的结果就是完全是由物理的力以及手柄和钢杆的材料特性决定。

表**4.1**　制造螺丝刀的步骤或指令

用电锯将 8 毫米粗的钢杆截为 10 厘米长。 用切削机以 11° 的角将杆的一端削（磨）至 3.5 毫米深。 沿杆轴翻转 180°，将杆头的另一面也以 11° 的角削至 3.5 毫米深。 用辊轧机将杆的圆头 12 毫米处压至 10 毫米宽。 将 25 毫米粗的塑料棒截为 10 厘米长。 在塑料棒一端 8 毫米处用车床车出 2 毫米深 16 毫米宽的拇指槽。 在塑料棒拇指槽一端的中心钻出 25 毫米深 8 毫米粗的孔。
在手柄上制造 5 条等距的 1 毫米深的槽。用 5 毫米钻头的固定钻，将手柄在转动的钻头上来回移动。
用锤子将钢杆的圆头用力插入手柄，让压扁处紧紧嵌入手柄。

表4.1中给出的每个步骤都表现出同样的基本模式。建立一个特定的状态，然后物理定律作用于状态就会产生特定的结果，结果由状态的

细节以及物理和化学定律预先决定。这9条指令步骤编码了一个掌握了必要工具知识的人以正确顺序配置必需的状态所需的信息。设计图包含了附加的信息，可以帮助制造者解读指令。也可以用语言描述图中所有的相关信息，但是没这么方便。重要的是，没有简单规则或公式能够预先确定或描述整个过程。在制造螺丝刀时，制造者总是遵循一个算法（复杂的规则）来实现所期望的结果。

表4.1中列出的9个步骤并不是全部。制造螺丝刀所需的信息还包括如何制造两个杆，制造所使用的各种工具和机器，以及操作机器的方法所需的所有信息。完整的指令表还必须包括采矿、炼钢、塑料生产和成型，制造所有的工具，以及制造这些工具的原材料。穷尽所需的指令会非常长；但原则上还是有可能列出制造螺丝刀所需的所有指令。

就算完成了指令表也还是不够。我们目前为止讨论的所有指令都需要由人来阅读、理解和完成。制造螺丝刀对人的要求包括：能够安排材料和使用工具的人体以及能够指挥身体、理解指令、解释设计图、理解任务和将理解转化为行动的大脑。现代分子遗传学已经发现了人体本身也是指令的产物。这些指令以DNA分子的化学结构的形式存储在我们身体的每个细胞中。现代神经科学还发现，阅读指令、解读设计图、理解和使用机器以及指挥身体的动作的内部信息编码在我们大脑的神经连接和通信模式中。指令、设计图、DNA序列和神经连接模式都编码了目的性信息。

现在我们能清楚认识到一块岩石和螺丝刀的区别。螺丝刀很特殊，是因为说明其形成所需的特定的低概率状态序列需要附加信息。只有发生特定的事件序列，我们才能制造出一把螺丝刀。岩石、盐晶、雪花、贝纳德流和太阳系的形成都不需要附加信息，因为它们形成所需的初始条件出现的概率足够高，到处可见。

没有钟表匠会有钟表吗？

想象一块老式怀表，里面有发条和很多齿轮的那种。许多零件的协同运动使得精确的时间呈现在表盘上。要让钟表工作，就要让机械部件正确匹配，要做到这一点，每个零件都必须符合严格的规范并正确安装。活细胞要比钟表更加复杂，但它们也需要准确构造的部件正确匹配

到一起。细胞的部件不是金属齿轮和发条，而是蛋白质，生命必需的大分子。细胞比钟表更复杂，部件也更多，功能更多样，但每个蛋白质都必须有正确执行功能所需的属性。

钟表匠的比喻有很长的历史。1802年神学家威廉·佩利记述了穿过一片灌木时捡到钟表的经历。他比较了钟表和旁边的岩石，他知道岩石是一直在那儿的。他想知道的是为什么钟表会出现在这里。如果认为钟表也像岩石一样从最开始就一直在这里，显然很荒谬，钟表必定有制造者，而且基于同样的逻辑，人类必然也是这样。理查德·道金斯在他的《盲眼钟表匠》一书中采用了这个著名的比喻 [2] 。书中道金斯认为创造生命的钟表匠其实是进化过程，而不是佩利所认为的那样！

要彻底理解道金斯的观点，我们需要了解细胞的运作原理，以及蛋白质的生成。在现代世界之前，所有蛋白质都是由活细胞生成。要认识生命就必须认识细胞。这些微小的结构单元有一些基本的特性，如表4.2所列。每一项都依赖于蛋白质的活动。

蛋白质通过与某物结合产生效用。同所有极微小的事物一样，它们是化学物质，并且同所有化学物质一样，它们的表面有原子尺度的特征，能够与其他物质原子尺度的特征产生吸引或排斥。这意味着如果某种蛋白质与另一种分子能够精确匹配并且表面的特征相互吸引，两者就可以黏合在一起。如果形状不匹配并且相互吸引的化学特征无法对齐，结合就不会发生。准确结合需要单个原子尺度上的空间位置能匹配，比钟表齿轮的啮合还要精密得多。细胞是怎样生成这种结构的呢？

表**4.2**　所有细胞都具有的**7**个特征

1. 由脂质蛋白组成的半流质膜隔开内外部分。
2. DNA 分子编码合成 RNA 和蛋白质的指令。
3. 具有合成蛋白质的机构。
4. 通过自我调节的代谢网络维持相对一致的内部环境。
5. 系统提供了所有所需部件的有序置换、修复和复制。
6. 细胞的生长和繁殖机制。
7. 获得化学能量和所需的原材料的机制。

在蛋白质的合成过程中，存储在DNA中的信息在细胞中扮演的角色同人类钟表匠执行的指令一样。用DNA写出的指令引导细胞中的机构准确组装特定的蛋白质分子。这个策略十分巧妙。许多步骤涉及自组装——免费的结构，但没有DNA中编码的指令，完整的过程就不会发生。

蛋白质是由名为氨基酸的小分子通过化学键连接到一起组成的链，就好像珠子串成的项链。氨基酸是相对简单的化学分子，有10到27个原子（比较一下，水分子有3个原子，乙醇有9个原子）。所有生物的所有细胞都基于同样的20种氨基酸，典型的蛋白质由数百个氨基酸组成，大蛋白质有数千个。

氨基酸都具有共同的可变结构部分，称为侧链。侧链的化学特性决定了氨基酸的特点。在一条链上氨基酸的相互作用（结合）会导致链向自身折叠，使得蛋白质在整体上具有特定的形状。这种形状及其表面的氨基酸的化学性质，决定了蛋白质可以与什么结合以及结合多紧密。图4.2展示了由129个氨基酸组成的还没有折叠的链，图4.3则是链折叠后自然形成的3维形状。

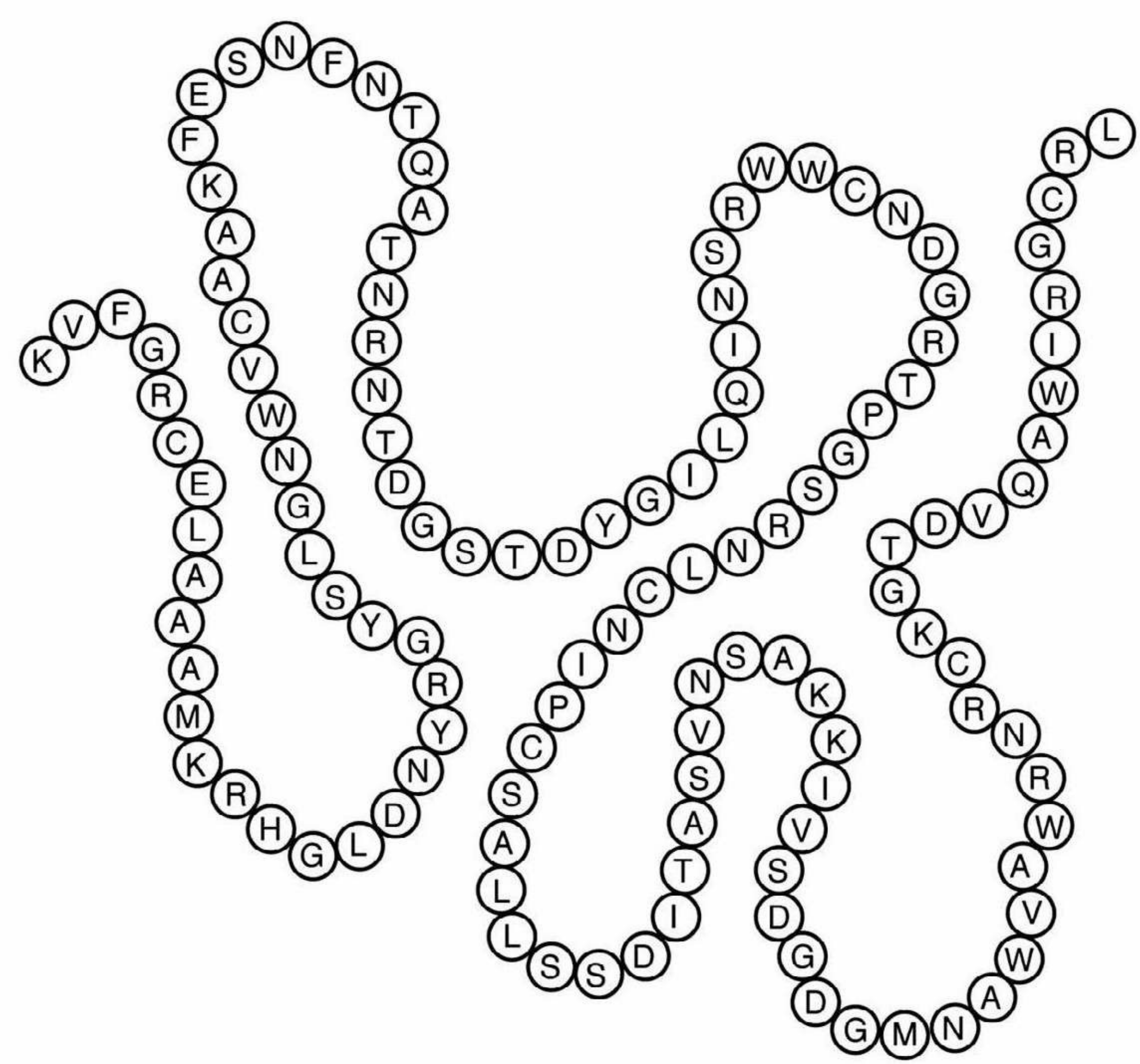
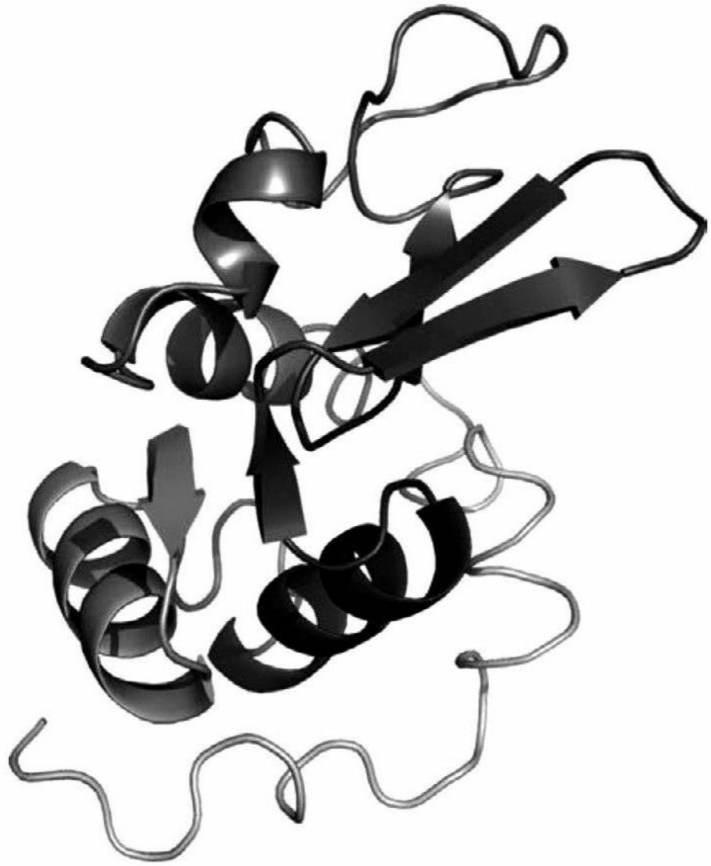


图4.2　129个氨基酸组成的溶菌酶的氨基酸链



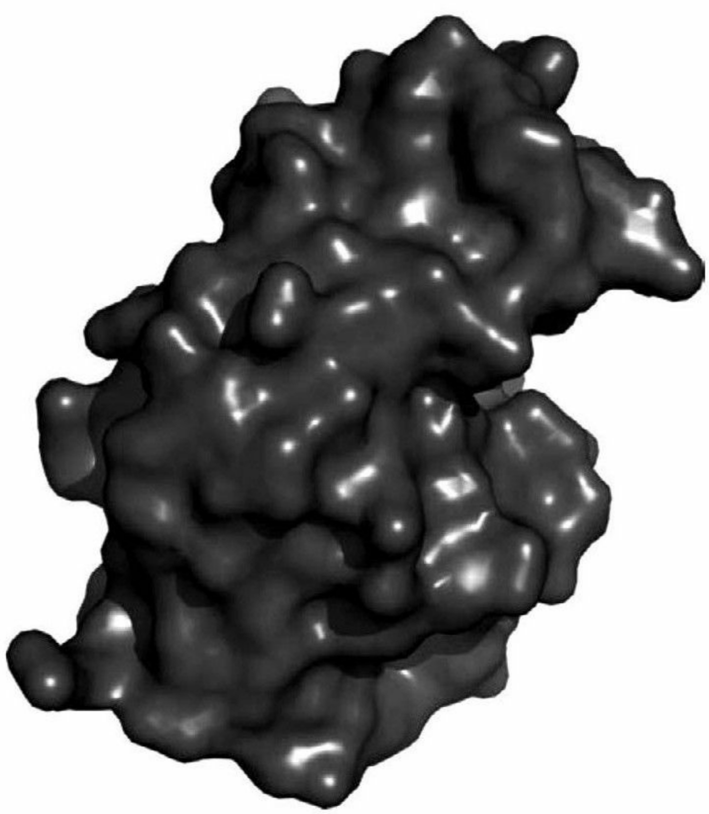


图4.3　溶菌酶的3维折叠结构图。带状模型（上图）展示了氨基酸链的构形。表面模型（下图）则展示了分子表面的所有原子。绘图软件采用薛定谔有限责任公司的PyMOL分子图像系统，版本号1.2r3pre

在自然界，所有蛋白质都以同样的方法合成。DNA中编码的信息决定每条蛋白质链上氨基酸的顺序，顺序又决定了形状。图4.2和4.3中展示了溶菌酶蛋白质都具有同样的氨基酸序列和同样的形状，暴露在表面的也是同样的侧链。就算只改变一个氨基酸，形状和表面的细节也会有所不同。

一些蛋白质催化（加速）小分子的化学反应。具有这种特性的蛋白质称为酶。酶的功能是与参与反应的小分子结合。一些蛋白质的功能是与DNA结合，还有一些是与特定的蛋白质结合。所有结合都很精确，并且蛋白质的氨基酸序列的任何变化都有可能改变其表面结构的细节，从而无法与预定目标结合，也许会与其他东西结合。人类细胞能生成数千种不同的蛋白质，都能以高度的特异性与细胞内部或外部的某种物质结合。

细胞的运转依赖于大量精确的蛋白质结合事件，因此正确的蛋白质合成对所有生命都很重要。这是通过微小的分子机制实现的，而这种机制本身又是依靠蛋白质和RNA, RNA是与DNA关系密切的大分子。RNA的合成需要蛋白质酶，蛋白质的合成又需要RNA。因此细胞合成蛋白质的机制也依赖蛋白质。

蛋白质合成是从DNA开始。与蛋白质类似，DNA分子也是由微小单元组成的链。蛋白质是由氨基酸组成，DNA分子则是由核苷酸组成。核苷酸有4种不同类型，一般简记为A、G、C和T。核苷酸可以以任意序列组链。因此，DNA分子可以用字序列描述，例如ACGATTCAAAGTCTCAG，其中每一个字母都代表DNA链中的一个核苷酸。这些字母携带了对细胞有意义的信息，就好像文字可能携带对你有意义的信息一样。细胞中的DNA分子很长，通常包含100万（$1·10^6$）到1亿（$1·10^8$）个核苷酸。核苷酸组成基因，每种基因编码一种蛋白质。

细胞中的DNA由两条核苷酸链相互缠绕连接而成，形成双螺旋结构。能够形成这种结构是因为A核苷酸可以与T结合，G核苷酸可以与C结合。这种配对使得双螺旋中的两根单链具有互补性：一边是A另一边

就是T，一边是G另一边就是C。这种冗余使得损伤可以修补，DNA分子也可以被复制。

合成蛋白质的第一步是复制DNA单链上的特定部位合成RNA。RNA的化学结构类似DNA，但使用的核苷酸有些不一样。它们也有单字母缩写，分别是A、G、C和U。一种名为RNA聚合酶的蛋白质负责合成RNA，合成的RNA的核苷酸链与选取的DNA段序列相同，不过DNA上为T的地方，RNA上为U（图4.4）。大多数复制的RNA是信使RNA（mRNA），但也有一些RNA具有其他功能。

mRNA分子可以与核糖体分子结合。核糖体的作用是将氨基酸按mRNA中核苷酸的顺序连接起来合成蛋白质。这样DNA中核苷酸的顺序就间接而准确地决定了蛋白质中氨基酸的顺序。RNA核苷酸只有4种而氨基酸有20种，因此并不是一种氨基酸对应一种核苷酸。mRNA链中3个连续的核苷酸组成的三联码对应蛋白质链中的一个氨基酸。4种核苷酸3个一组总共有64种排列（AAU, ACU, GAU，……）；每种三联码称为一个密码子。64种密码子中有61种都对应特定的氨基酸。还有3种是“终止”密码子。这个合成过程称为转录，意指核苷酸编码的信息被“转译”成氨基酸编码的信息。

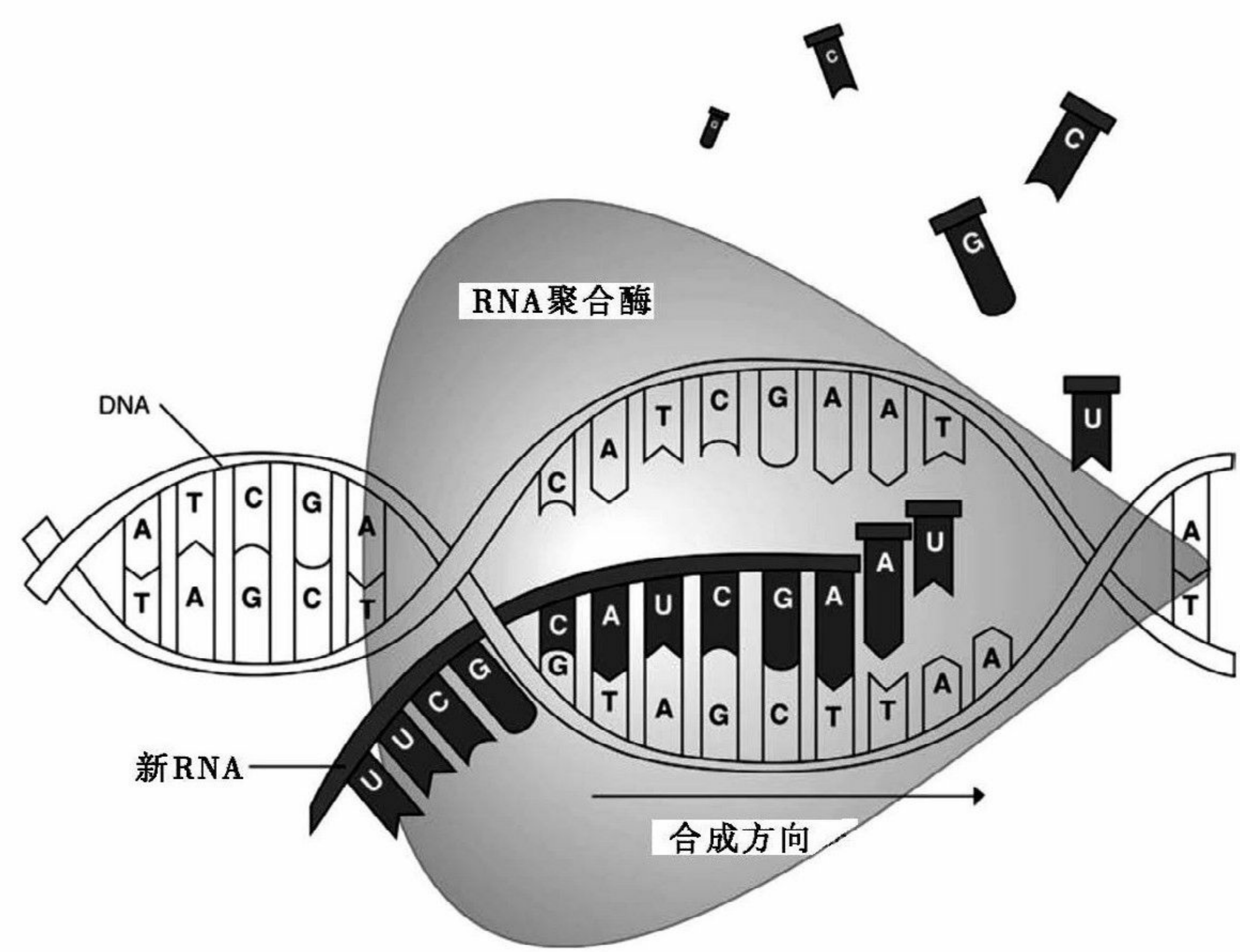


图4.4　RNA聚合酶从DNA序列复制RNA的示意图

氨基酸与核苷酸的结合效率不高，还需要名为tRNA的适配器分子在核糖体中将正确的氨基酸与对应的mRNA密码子对齐（图4.5）。这种适配器确保氨基酸与每3个核苷酸对齐。这个过程的正确性很重要。一旦出错，生成的蛋白质就会有错误的氨基酸序列，从而可能折叠成错误的形状并且无法正确执行功能[3]。

越来越多的氨基酸被合成后，会从核糖体伸出来，氨基酸相互作用，使得尚未成形的蛋白质开始折叠成由序列决定的3维形状。大多数蛋白质在合成结束时就已经成形了；有一些则需要后续修饰，还有一些要等到与目标物结合时才会最终成形。

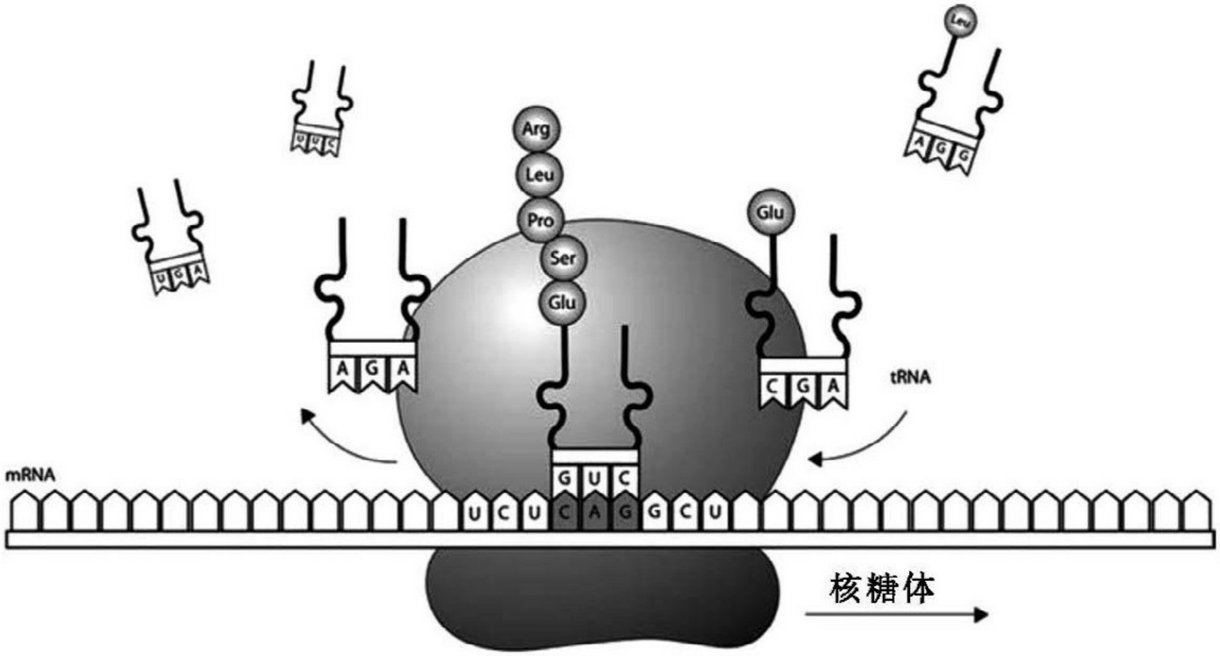


图4.5　核糖体、mRNA和tRNA适配器参与蛋白质的合成。深色结构是与mRNA（长链）作用的核糖体，tRNA与mRNA碱基配对，同时还连着一个氨基酸链。图中下一个要结合的tRNA正在从右边进入核糖体，前一个tRNA则正从左边离开。中间的tRNA在与mRNA的3个核苷酸配对后，连在它上面的氨基酸链将被转给后面进来的tRNA

自发折叠是理解氨基酸链编码的信息如何转换成3维结构的关键。氨基酸会相互作用，因此链会寻找优化所有相互作用的总能量的形状。过程的这个部分是免费的结构。存在大量可能的蛋白质形状，因为有大量可能的氨基酸序列，每种都会折叠成不同的形状。这就是蛋白质的生成过程，细胞的正确运作需要准确生成的蛋白质。每个蛋白质的生成都遵循编码为DNA的指令。正是DNA序列确保了正确的蛋白质被生成出来。改变序列，就会改变蛋白质。生物不是钟表，而且要复杂得多，但它们的出现无需钟表匠的介入。

DNA编码的指令如何说明比单个蛋白质更复杂的结构？

我们继续钟表的话题，不过现在想象一种机械装置，不是由钟表匠组装，而是将零件放在箱子里摇晃就能自行组装。滑稽吗？能不能设计出这种零件放在一起摇晃就能自行组装的钟表？需要什么样的零件？当然我们所知的钟表零件做不到，但病毒的确是这样形成的。

病毒是像钟表一样复杂的生物装置，但要小得多。病毒的分子组装

就类似于在箱子里摇零件。这是如何做到的？首先，零件都很小，一旦悬浮在水中就会由于热运动自发跳动，这是所有极微小事物的特点。第二，零件是蛋白质，我们已经看到，所有蛋白质都有很精密的形状。第三，它们的表面具有能与其他病毒蛋白相互作用的分子特征。一旦适配，蛋白质之间就会紧密结合。

病毒能够繁殖，但并不生长，也没有独立于宿主细胞的化学代谢。在存在的大部分时间里它们都完全不活动，介于生命和非生命之间。病毒的繁殖很像塑料玩具的组装，零件的凸起与其他零件的凹槽相嵌。区别在于没有爸爸或妈妈来组装。同其他蛋白质一样，所有病毒蛋白质的形状都是由特定的氨基酸序列决定。形状（以及相应的表面特性）决定了它能结合什么样的形状。将一个病毒蛋白质放到正在成形的病毒的正确位置上所需的所有信息都编码在其结构的细节里。

可能没有哪种活细胞能抵抗一切病毒的感染。幸运的是，病毒具有高度的特异性，任何特定的病毒都只能感染有限种类的细胞。感染细菌的病毒尤其容易研究。它们被称为抗菌素或噬菌体。名为T_4 的抗菌素是被研究得最多的细菌病毒之一。它完全由蛋白质和DNA组成。这种病毒相对较大也较复杂，很适合用来说明自组装的潜力。T4的DNA编码了274个基因，分别对应病毒组装和繁殖所需的274种不同的蛋白质。其中一些蛋白质是改变受感染细菌细胞的代谢的酶；一些则是合成T4的DNA所需的酶；还有一些参与274个基因的表达。大约有50种蛋白质用于建造其怪异的外壳，当病毒从一个濒死的细胞中释放出来时，外壳可以保护其DNA。这个外壳还用于在遇到下一个适合的细菌细胞时对其进行感染。T_4 的外壳由40种3500多个蛋白质分子组成。图4.6画出了它的结构。DNA隐藏其中，缠绕在“头部”。幸运的是这种病毒不会感染人类细胞。

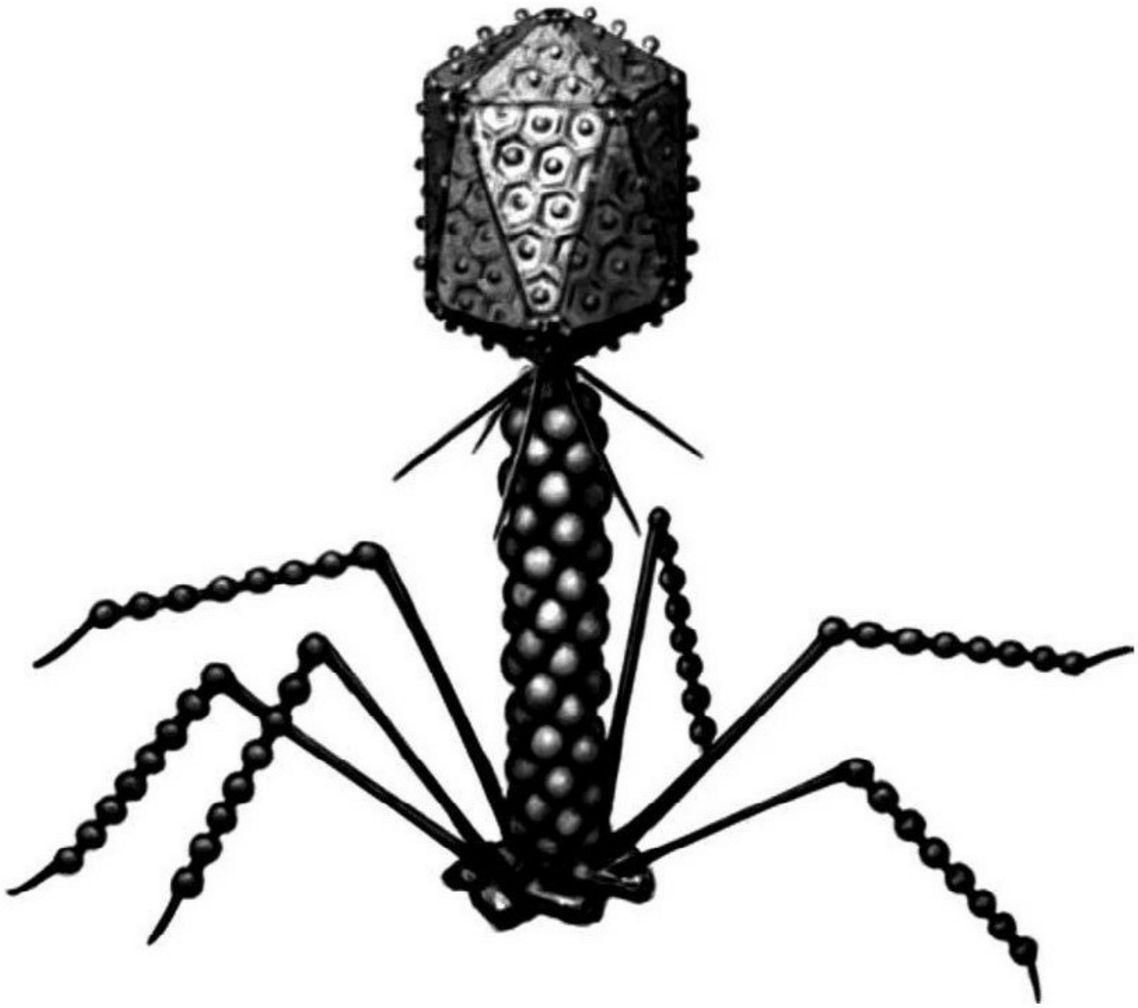


图4.6　T4噬菌体（参见注释4）

与其怪异的结构相比，它的形成更让人吃惊。并没有什么小精灵围着它用微小的锤子将所有零件装到一起。这个结构是自组装的。其中的3500个蛋白质都只能与特定的伙伴分子匹配。这本身倒没什么；毕竟，汽车零件就是以特定的方式组装到一起。但汽车无法自组装。汽车零件的设计并不是让你可以将它们放到一个大箱子里摇晃，然后就能组成一辆车。蛋白质很小，受周围快速运动的水分子影响，会不断跳动。当相互匹配的形状发生接触，就会结合到一起。每一次结合都会产生新的结合点，可以让另一个蛋白质结合。这建立了一个组装顺序。只有前两个蛋白质结合到一起，第三个蛋白质才能结合，然后第四个才能结合，依次进行。这样蛋白质就以特定的顺序被添加到生长的结构上；从而从随机混合的蛋白质溶液中就能自发出现特定的结构。

这种组装通过3条并行的“组装生产线”进行（图4.7），头、尾和长尾纤分别组装，最后才组装到一起形成完整的噬菌体。在头部与尾部相连之前，它会利用被感染细胞提供的能量装入T_4 的DNA。

这个过程很像图2.10中的自组装拼图游戏，但结构是3维的，图块是蛋白质。拼图和病毒的形成都有赖于让形状能以预先确定的方式精确组合的目的性信息。对于拼图，设计图块的信息是由人的大脑创造。对于T4，信息编码在病毒的DNA中。如果没有说明这些组块细节的指令，病毒（或自组装拼图）就永远都不可能出现。

自组装病毒与第2章讨论的分子砖块以及新科学（NKS）系统都属于同一类活动。这类系统通过组装计算物理图样和结构。计算的规则——程序——包含在组装部件的形状和化学结合特性中。

细胞中的许多活动也是基于自组装蛋白质结构的活动。实际上，整个细胞甚至整个躯体都可以视为大量自组装事件的产物。小部件以特定的方式组成大部件，反过来又成为更大更复杂结构的基础。在已存在的结构之上建立预先确定的结构。从这个意义上可以说细胞就是活的计算机。细胞中几乎所有的结构最终都依赖于用DNA语言编码的指令，DNA决定了蛋白质和RNA，它们的相互作用以及它们的形成。

指令能引导活动吗？

由于病毒很小，在水中随机跳动。给以时间和运气，这种随机运动最终会将病毒带到合适的宿主细菌。在6条长尾纤的顶部（图4.6）有结合点，能与宿主细胞表面特定的分子结构匹配。一旦某条长尾纤与目标结合，很快其他长尾纤就也会相继与细菌表面的目标点结合。这时，T4就会头朝上尾朝下稳稳地附在细菌上。长尾纤的成功结合会引发蛋白质的变形。这个过程很像变形金刚玩具，零件以特定的顺序移动使得车或飞船变成机器人。变形金刚可以恢复原状——病毒不会。

一旦长尾纤结合，病毒底部名为短尾纤的蛋白质就会展开，与细胞表面的其他特征结合。这又会导致基座从六边形变成星形，触发144个尾鞘蛋白质展开。这会使得尾鞘缩短大约2/3，驱动内尾管像注射器一样穿过细胞外膜。宿主细菌细胞有3层防线，外膜、细胞壁、和内膜。

T4的DNA要成功进入细胞，就必须离开噬菌体的头部，通过尾管穿过细胞的3层防线，进入细菌的细胞质。尾鞘收缩会将尾管顶端顶在第2层细胞壁上。这是一层坚固的结构。管端部的酶会解开细胞壁上的化学键，让其变弱，从而使得管子可以穿透进去。这样尾管端部就顶到了细胞内膜。然后管蛋白与内膜蛋白相互作用，形成一个孔，让DNA得以穿越最后一道屏障，进入细胞质，在这里开始合成噬菌体蛋白[4]。

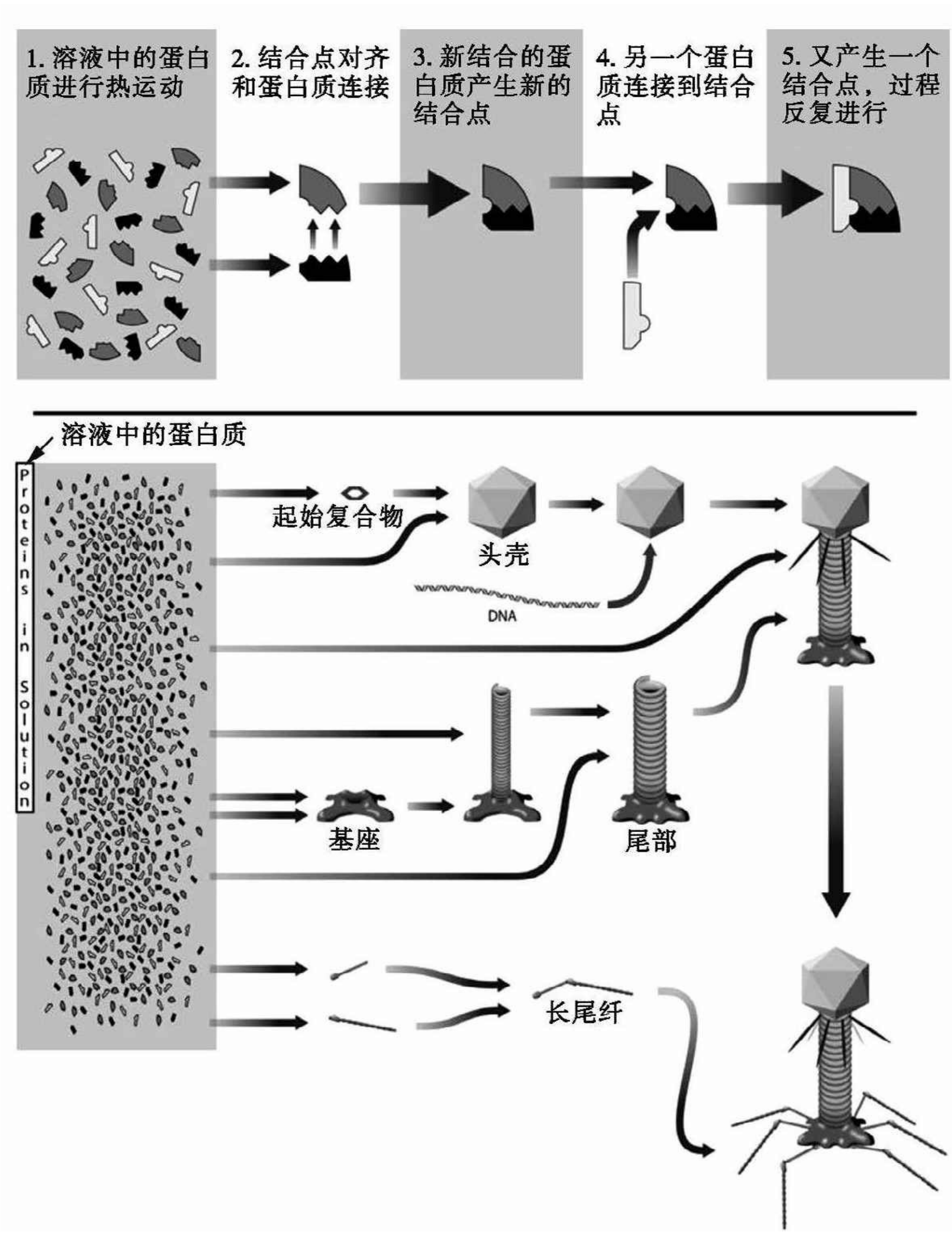


图4.7　T4噬菌体的组装（参见注释4）

在结合和注射的生命阶段，T4噬菌体短暂变成非常活跃的结构，大量蛋白质变形，并且有两次大规模重排，尾鞘收缩和DNA注射。T4不能生产代谢化学能，因此势能在组装的时候就储存在其结构中。就好像发条玩具被放开后在地板上移动。一旦储存的能量被消耗掉，病毒外壳就无法再活动。与玩具不同的是，这里没有可以重新上紧的发条。T_4的活动完全取决于其结构以及蛋白质与环境的相互作用。感染过程中发生的具体动作都是由T4的DNA编码的信息预先决定的。特定的外部信号触发预先设定的动作。我们可以认为这些动作本身是预先组装好的，等待触发。

持续进行的活动呢？

噬菌体感染是很短暂的顺序执行的动作，由病毒结构预先决定，以达到低能态。下面我们想知道的是长期持续的活动是不是也能预先编程。这个问题的答案当然也是肯定的。我们的例子同样来自生物学。有一种基于细胞的系统能不断产生活动波。这种波被称为动作电位，由交替产生的离子流造成，它是我们身体里的神经细胞相互通信的基础。这种波的模式形成了一种时间上的非平衡态结构，它们需要来自细胞代谢的持续能量输入。回忆一下，离子是带电原子或分子。这个例子也为我们后面讨论人类学习建立了分子基础。

身体里的细胞分别执行不同的任务。肌肉细胞专门产生物理力，血红细胞传输氧，白细胞则负责侦测和消灭外来细胞和病毒。神经细胞的任务是在身体里进行远距离通信。细胞的特性和活动主要是由它们的蛋白质决定，而就如我们知道的，蛋白质又是由DNA序列（基因）表达为mRNA分子，然后又转录为氨基酸链，并且折叠为具有高度特异性的结构形成的。

所有细胞都维持一个不同于外界的内部环境。其中一个不同之处就是离子的浓度。尤其让人感兴趣的是钠离子和钾离子。钠离子就是食盐里有的离子，在身体里，主要分布在细胞外部的血液和间质液中，而钾离子则是细胞内部的浓度更高。在细胞膜上有一种被称为“泵”的特殊蛋白质不断用细胞内部的钠离子交换外部的钾离子来维持这种差别。它们通过在两种结构状态之间不断来回转换做到这一点。每次状态转换都会携带离子穿过细胞膜。一次翻转将钠离子送出去，下一次翻转将钾离子

带进来。这种泵送活动需要能量，能量由细胞内部其他蛋白质酶参与的化学反应提供。

许多神经细胞（神经元）都很长很细。你身体里最长的神经细胞从下脊椎一直延伸到脚趾尖。长颈鹿和巨型乌贼的可以有几米长。信号通过移动的动作电位沿这些细胞传递。动作电位由短暂的钠离子内流和紧接着的钾离子外流组成。这个模式沿着细胞膜前进。这个短暂的离子流会引起电脉冲。从信息的角度看，动作电位可以视为传递了一个比特的信息。复杂的动作电位模式可以编码更复杂的信息。

在分子层面，动作电位并不复杂。它是细胞膜上称为“通道”的微小蛋白质门的顺序开启和关闭所引发的。这些通道的结构与钠/钾离子泵蛋白质不同，它们可以选择性地让离子通过，无需消耗额外的能量。有两种通道与我们的问题有关，一种是钠离子通道，还有一种是钾离子通道。自然有消除化学浓度差别的趋势，因此如果两种通道都长期开启，细胞内部和外部的钠离子和钾离子浓度就会相同，细胞会死亡。由于门只是短暂开启，因此细胞不会受到损害。

选择性的离子泵和离子通道的组合在神经细胞的细胞膜两侧产生微小的电位。这被称为“静息电位。”通常介于-60到-70毫伏（mV）之间。毫伏是千分之一伏。神经元离子通道的开关受电压影响，所谓的“电压门控”。一旦细胞的静息电位与0的差低于某个阈值，通常约为-50mV，就会引发动作电位。这种电位变化一般是通过与其他神经细胞的互动产生，但科学家也可以诱使其发生。

一旦达到阈值，电压门控钠离子通道就会通过相继的3个结构状态发生一系列变形。关闭的通道对电压敏感的状态称为静息状态。一旦膜电位减少到-50 mV，通道就会打开，允许钠离子进入细胞。这个状态大约持续1毫秒（千分之一秒）。然后通道自行关闭，进入短暂的对电压不敏感的状态。大约2毫秒之后，蛋白质重新恢复原来的构形。钠离子通道的整个循环花费大约3毫秒。第二个状态的短暂延时对动作电位的形成很关键。

钠离子的内流会局部改变离子平衡，使得细胞膜的一小片区域的电位短暂地从-50mV变到+30mV。这个变化有两个效应。它会触发附近的电压门控钾离子通道开启，并且触发稍微远一点的钠离子通道准备开

启。等到最先开启的钠离子通道开始关闭时（开启之后1毫秒），旁边的钾离子通道正在打开。这使得细胞内部局部区域的钾离子外流（带走携带的正电荷），使得局部的膜电位恢复正常。等钾离子通道完全打开时，旁边的钠离子通道正在关闭，而稍远一点的钠离子通道则开始打开。然后，钾离子通道又跟着钠离子通道打开。结果是钠离子内流紧跟着钾离子外流交替进行。离子流模式沿细胞表面移动，产生出短暂的电压变化波。这个过程如图4.8所示。

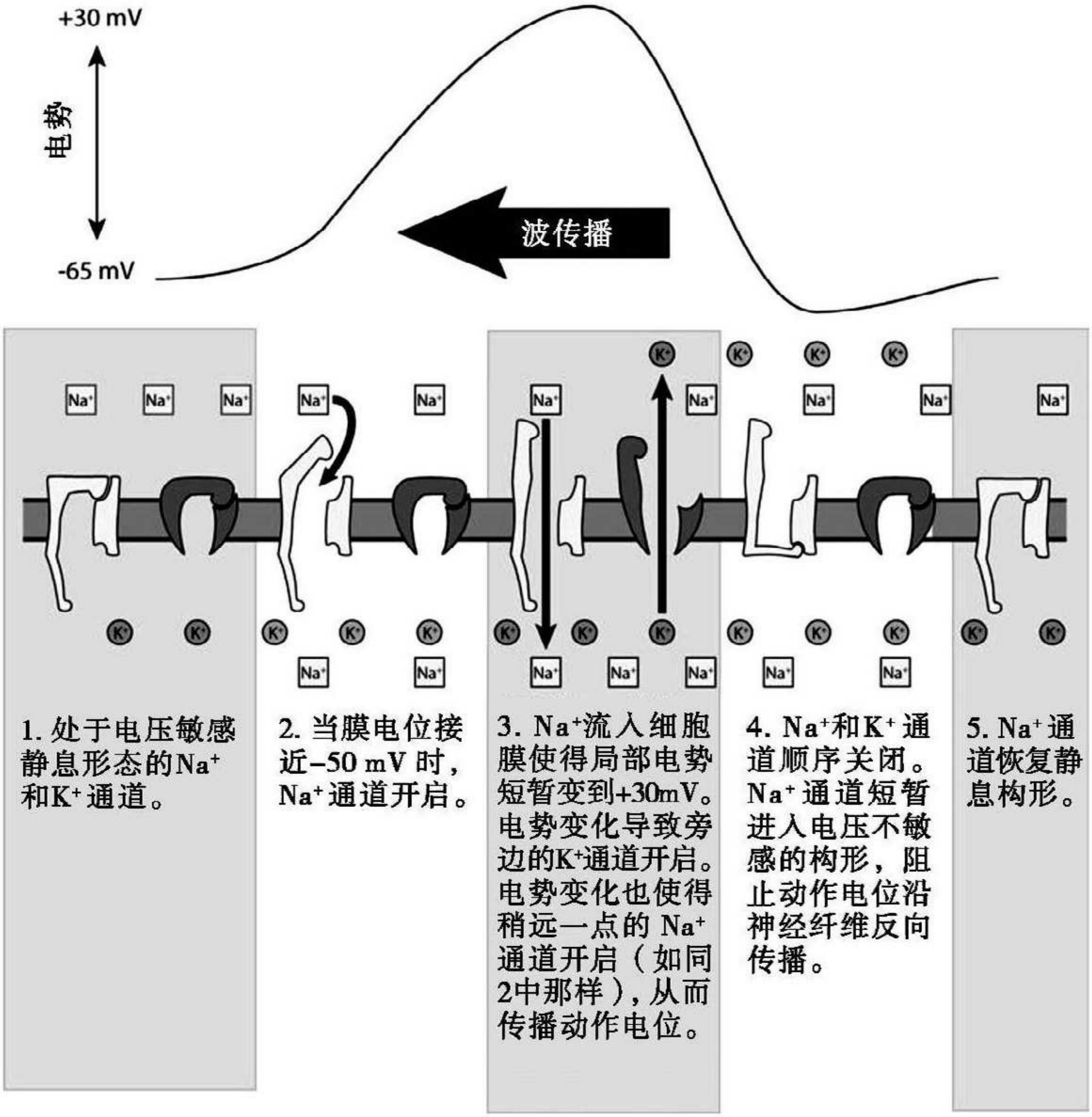


图4.8　细胞膜的离子流引起电压变化，形成动作电位。电压波从右往左传递。（参见注释5）

从钠离子内流到钾离子外流，然后恢复正常，整个循环大约5毫秒。动作电位沿细胞表面传播。图4.8中展示了细胞局部的电势变化。钠离子流入细胞导致电势上升（从负往正变化），然后接着钾离子外流又导致电势下降。

由于电压门控钠离子通道的变形周期需要几个毫秒才能完成，使得波（钠离子内流跟着钾离子外流）只能从激活的初始位置往外传播。在

波通过后，钠离子通道会短暂处于对电压不敏感的状态。这个电压不敏感时期确保了动作电位只能单向前进。

如果神经细胞是球形并且足够大，则波会绕着细胞一圈又一圈传播，但神经元又长又细，因此只能沿着细胞传播直到端部[5]。如果神经元分叉，动作电位会前进到每个分叉顶端，不会衰减。在顶端是称为突触的特殊结构，与其他神经元相接触，使得一个细胞的动作电位可以改变接触的神经元的静息电位。突触对其他神经元的作用可以是兴奋或抑制。神经元整合同一时间到达的所有信号，如果静息电位达到了阈值，就触发新的动作电位。这样动作电位就能选择性地在神经元之间传递。如果细胞的负输入超过正输入，输入的信号就无法传递。这使得系统可以处理信息。

神经元细胞膜上的蛋白质具有非常特殊的性质，它们的协同动作产生的动作电位组成了动态可更新的模式。关键的蛋白质是维持静息电位的钠/钾离子泵和电压门控钠离子和钾离子通道，当它们被触发就会顺序产生离子流。这些蛋白质的特性，尤其是电压门控钠离子通道的周期变形和这些通道在关闭前保持开启的特征时间，使得只要细胞膜上具有这些特殊的蛋白质，并且膜电位上升到阈值，动作电位就必然发生。由于动作电位是特定蛋白质的活动引起，而这些蛋白质的结构又是由DNA序列决定，因此可以认为动作电位是由指令预先决定的。同噬菌体感染一样，动态活动最终来源于DNA序列中编码的用于建造特定蛋白质结构的指令。

那么到底如何定义目的性结构？

前面在第2和第3章给出的例子表明，在静态和动态系统中，没有指令也可能出现显著的复杂性。这一章给出的各种复杂事物的例子则是只有通过使用指令才有可能出现。区别是什么？从多个角度来看，雷暴都要比动作电位更复杂，当然也比螺丝刀复杂。然而，从某些方面来说，雷暴却更简单，当然也具有更多自发形成的可能性——概率大得多，并且除非你相信雨神，否则雷暴不会被认为具有目的性。

螺丝刀、病毒、膜蛋白和动作电位有3点不同于波浪、太阳系和雷暴：预先设定、目的以及极低的概率。这都可以用指令解释。指令本质

上就具有预先设定某种事物的目的性，并且设定信息的量决定了，如果不使用指令，事物有多不可能。自然定律也是一种设定，但是它们简单、普适，并且不针对特定的结果。比较起来，指令更复杂，没有普适性，并且针对特定的结果。指令总是具有某种动机；它们代表对结果的投资。化学和物理定律简单作用于有合理可能性的初始条件产生出的事物可能会让人惊奇，也可能很精巧或很宏大，但如果我们排除宗教解释，就不能说这种事物是为了某种目的而设计，或是被预先确定适应某种东西。

螺丝刀和病毒无法仅仅用简单规则作用于常见的初始条件来解释。必须使用有针对性的特殊规则（指令）才能建立起创造某种本来出现概率极低的某物所必需的条件。当结构是作为使用指令的结果出现，产生这种结构所需的指令的低概率性就直接解释了这种结构的低概率性。既然随机生成哪怕很短的指令的概率都极低，我们就能肯定它们所包含的信息不是随机组合的比特。自组织的思想也无法回避这个结论。远离平衡态现象学为我们打开了动力学王国的无限可能，但并没解决低概率问题，也无法解释目的性。这个只有指令能做到。

如果我们暂时不考虑人类的创造性，唯一所知的能创造目的性信息（当然包括指令）的策略，就是可复制对象组成的群体，通过反复的累积性选择，自然执行的概率计算。我称这种计算策略为复杂引擎。正是这种计算策略支撑了生物进化和我们在后面的章节中将讨论的一些现象。如果没有累积性选择，就不可能在宇宙的时间和空间里找到编码特定结果所需的目的性指令的特定符号组合。选择是非随机的，选择规则为从无穷多的无用选项中识别出有用的信息组合提供了指南针。甚至当被选择的信息是随机生成时这也一样成立。目的性结构之所以存在是因为它们通过了具有某种目的性的选择；它们符合标准，标准决定它们继续存在下去要比那些没有被选择的更好。基于指令的对象和行为之所以具有目的性，是因为它们的指令是由在反复进行的选择中起作用的规则塑造的。

这并不是说波浪和雷暴的存在就没有原因；只是说其原因很直接。它们是对常规条件的自发反应。波浪、雷暴和沙堆崩塌都是传递能量的机制。要理解它们，我们需要知道在明确定义的系统中运作的物理力以及对这些力的阻碍。阻碍能量流导致势能增加，有时候表现为某种结构。其余的就只是细节。在我们的宇宙中免费的结构很重要，但仅此并

不能解释我们的存在。我们还需要信息的概念，适当组织的信息能引导概率极低的目的性对象的形成。没有这种额外的信息，自然就只能限于展现无穷的可能结构中概率最大的那些。然而，通过利用适当的指令，任何逻辑上可能的结构都能实现——无论其概率看上去有多低（假设有所需的资源）。这是十分惊人的论断。

目的意味着需要。创造某种事物是为了今后满足特定的标准。这个概念对贝纳德流没有意义。贝纳德流的形成是因为分子的协同运动正好能最有效地传送能量；这其中并没有事先的计划。只要一盆水被放在火上，就会出现温度差，而从水的底部往上部传送热的最佳机制就是对流运动。一旦几何条件正确，对流就会表现出规则模式。决定热传输的定律并不是专为形成贝纳德流而创造。

病毒将机械能和化学能转化为热能，从而参与通过生物圈的能量流，并最终参与从太阳到外太空的能量流；但病毒并不是做这件事最有效率的机制。它们的存在的重点不在于能量传递，而在于结构的延续——这一点相当不同。病毒结构的动机不在于释放能势，而在于延续。这种驱动力利用能势，但并不以其为动机。蛋白质也可以以同样的方式理解。蛋白质组成的更复杂的结构——我们称为细胞——也是受延续驱使。蛋白质可以被认为具有让细胞、病毒或生命生存的目的；而所有生命都有让它们的遗传信息延续的目的。有无数多种可能的蛋白质；实际存在的那些之所以存在是因为它们对生物体的延续有贡献。要了解基因延续的驱动力，可以阅读道金斯的《自私的基因》[6]。延续的驱动力是复杂引擎运作的核心特征和自然结果。

其他事情也都是一样的，我们预期所遇到的结构只要有合适的条件和历史就具有合理的可能性。这就是为什么来自其他行星的人会将螺丝刀视为智慧生命的象征。螺丝刀不可能存在，除非这个世界上有人用螺丝来造东西。它们存在是因为人类有拧螺丝的需要（目的）。同样，螺丝的存在又是为了将事物固定到一起的目的。这个动机的链条可以很长，但是我们如果想理解目的性复杂对象就无法回避它。太阳系和山脉也是产生自很长的事件链条，但那是受物理力驱使的链条，而不是需要。挑战在于理解需要如何转为目的，目的又如何转为满足需要的结构。这是下一章的主题。

第5章　无概率性和复杂引擎

有序和无序有何关联？

见过以前老式的电视机吗？我还有一台。当调到没有台的频道时，屏幕就会显示随机闪现的“雪花”点。换一个有信号的频道，会看到演员表演或很远地方的风景。屏幕上有序和无序的区别很明显。雪花屏幕与有人说话的屏幕的像素是一样多的，但它们的状态不一样。盐晶、病毒和动作电位是自然界中有序的例子；气体和液体则明显缺乏结构。如果组分（像素或分子）之间有关联，我们就会识别出结构。结构体现有序，虽然根据我们的经验有序和结构很常见，它们的定义却比你可能认为的要更难琢磨。

抛硬币是一个很好的例子。如果反复抛，预期有50%是正面，我们说某次为正面的机会是50/50。有许多手法可以改变预期的结果，但只要你使用的硬币匀称，抛1000次大致是500次正面500次反面。显然如果你在硬币正面标H反面标T，抛1000次硬币会得到1000个H和T组成的字符串，这个字符串将是一个“随机”无结构的1维字母排列。但真是这样吗？随机生成就能保证随机序？

回答是否。如果抛硬币实验反复进行许多次，结果偶尔也会表现出明显的模式。如果你抛了一次正面，下次反面的机会是50/50（概率1/2）。如果你先抛了一次正面，又抛了一次反面，下次为正面的机会还是50/50。因此依次抛出正、反、正（H·T·H）的概率为1/2×1/2×1/2，即1/8。抛出H·T·H·T·H·T·H·T·H·T的概率是1/1024。原则上，如果抛1000次硬币，以下结果都有可能出现：1000次交替的正和反，2次正跟着3次反（H·H·T·T·T）重复500次，500次正跟着500次反，甚至1000次反，1次正都没有出现。虽然这些有序的结果理论上是可能的，我们却预计永远也不会看到它们发生。这是因为抛1000次硬币得到特定结果的概率是$1/2^{1000}$，小得难以想象。如果预先选定某个特定的结果，我们一直抛到宇宙终结也无法看到这个结果发生。

抛硬币揭示了一个对所有由组分组成的对象（或系统）都适用的原

理：组分的随机排列通常会导致无序，但并不总是这样。这个原理对符号序列、电视机屏幕上的像素以及分子的3维构形都适用。分析表明组分的数量越多，非随机状态在可能状态中的占比就越低。极端条件下，例如2个字符长的H和T排列似乎都有序，要么是相同符号出现两次，要么是交替出现。在另一个极端，无穷长序列中有序的可能性为0。

对这类讨论的一个挑战是对有序的认识有主观性。幸运的是，有严格的方法可以定义随机和非随机，这个方法来自计算理论。根据这个思想，随机序列的元素之间没有内在关联，即没有相关性。我们在第2章讨论过这个思想。例如，假设元素是英文字母，那么对于随机的字母序列，没有什么方法能预测序列中某个特定的字母，就算其他字母都给定了也是如此。而非随机序列则具有内在关联，可以用序列不同部分之间的关联规则来刻画。例如1000个A后面跟着1000个B。存在简单规则能将A和B关联起来。规则“写1000个A接着1000个B”就能得到想要的序列。

在计算机科学中，计算的输入被视为对输出的描述，而有序和无序对象的定义区别是有序对象允许短描述。对于很短的序列，例如H·T，存在关联元素的简单规则，但规则比序列本身更长也更复杂；因此我们无法说这种序列是不是有序。对于H·T·H·T·H·T·H·T·H·T，规则为“重复H·T　5次”，如果我们缩写为5×HT，会比原来的序列短一点。对于1000个交替的H和T，规则“1000×HT”就要短得多，因此我们可以明显认识到这个序列是有序的。

计算有一个不那么容易认识到的特点是，大多数输出都无法用比输出短的输入得到。绝大多数的长序列都是随机的。如果输出存在内在关联，输入就有可能比输出短。“有趣的”计算对象几乎都是那些具有内在关联——即具有结构——的形态。有序的呈现有时候很微妙，很难通过眼睛发现，但如果一个对象（输出）可以用短输入通过确定性计算得出，我们就知道其中存在某种形式的有序。这为有序提供了定量定义：一个对象如果存在比对象本身短（或小）的描述，就具有有序性。

关于对象的随机性还有很重要的一点是有序和随机不是非此即彼的关系。一个对象可以很有序或不那么有序。事实上，很多对象都是既有序又无序。抛一个不匀称的硬币生成的H和T长序列可能是55%的T和45%的H。这样的序列不是非常有序，但还是表现出一定的规律性。我

们的身体有许多结构，但在细胞和亚细胞层面上也有许多随机性。

总结一下，有序排列的特点是存在内在关联，而所有物理系统都共有的一个重要特征是绝大多数可能的形态都没有内在关联。对于大系统，有序不符合预期，因此需要解释。

如何弥补现实极端的不可能性？

道格拉斯·亚当斯的《银河系搭车客指南》中有一段有趣的话：

无穷小概率驱动是一种惊人的新方法，可以在几乎不到一秒钟的时间穿越星际间遥远的距离，不再需要在超空间中单调乏味地东碰西撞。当小概率驱动达到无穷小概率时，它会几乎同时穿过所有可想象的宇宙中所有可想象的点。也就是说，你永远不知道你会到达哪里，甚至不知道你到达时是什么物种。这对于穿什么衣服很重要[1]。

在亚当斯的有趣故事中，没有穿太空服的阿瑟·邓特被抛入深空，在必死的情况下被黄金之心救下，黄金之心是一架神奇的飞船，利用“无穷小概率驱动”在太空中旅行。阿瑟·邓特获救的无概率性是“2的27.6万次幂分之一”（$1/2^{276000}$）。任何读者都会认为这很荒谬，但为什么呢？要回答这个问题，我们需要探讨一下当我们说某件事情有存在的概率时是什么意思。

想象一个空间，空间中的每个点代表一种可能存在的事物。真实存在的事物在其中只占极小的一部分。在这个空间中的移动就是由一种可能变成另一种可能。变化如果没有规则，就只能向随机选择的新位置跳跃，很像黄金之心。问题是在这个想象的空间中绝大多数可能存在的对象都是随机的。因为它太大了，像你我、太阳系和摩天大楼这样有序的事物在这个空间中占的比例微乎其微。如果黄金之心在由所有可能性组成的空间中跳到一个随机选择的位置，这个位置不会是一个有序的状态，只会是完全无序的状态。

结构之所以存在是因为宇宙根据非随机定律运作，并且有特定的历史。没有规则，就不可能存在有序。在物理宇宙中会出现结构，是因为化学和物理定律限定了可能的结果，结果衍生于之前的结果，而之前的结果本身就具有结构。一方面，物理定律“解释”了为什么宇宙会表现出

结构，另一方面，这些规则的来源又是关于我们的存在最深的迷团。重要的是要记住，这些规律在让有序成为可能的同时，并没有禁止随机性；毕竟，抛硬币的结果就是随机的。

400年来科学家观察宇宙，推导运作规律，并用实验验证他们得出的规律。这些规律决定了在可能性空间中什么样的路径是允许的，因为它们禁止了许多事件，不可能把黑洞变成银币或把银币变成阿瑟·邓特。允许的事件又为其他允许的事件创造了条件，串接起允许的路径。在允许的路径中，有很多时候允许发生随机行为，有时候随机才是常态。举个例子，物理学家知道铀235原子会衰变成钍231和阿尔法粒子，但他们不知道衰变会在什么时候发生。一个特定的原子在未来7亿年里衰变的概率是50/50。一个特定的铀235原子在我们活着的时候衰变的可能性很低，但不是0。一小块铀就具有放射性，因为它含有如此多的原子，以至于每一秒都有衰变发生。（原子）变化的时间很随机。这个现象可以用方程描述，但只是作为大量原子的统计特性或衰变的概率。

在所有可能性组成的空间中，随时间变化的路径有一个特性，未来的状态根据规律衍生自过去的事件。结构基于过去的结构。还有一个与放射性衰变有关的例子，磷32是磷元素不稳定的同位素。磷是DNA的组成部分，并且常见的同位素磷31很稳定。有些磷32原子碰巧会被合成到DNA中。一旦DNA中的磷32原子衰变，DNA链就有可能断开。如果磷32原子所在的基因编码合成黑色素（赋予哺乳动物皮肤和头发黑色的色素）所需的酶，只要它不衰变和破坏DNA就不会有影响。如果衰变发生在精子或卵子的细胞核中，生成的胚胎就有可能有黑色素合成缺陷。通常这种突变是隐性的，不会马上表现，但如果未来某个后代将两个突变基因组合到了一起，这个个体就会白化。如果这个生物是野兔，就很可能会被猎杀，因为除非地上有雪，否则它很容易被捕食者发现。如果这个生物是生活在原始社会的人类，他或她也许会被认为具有魔力。另一方面，如果发生衰变的精子没有参与繁殖，也就完全不会有生物学影响。没有影响、被吃掉或具有尊贵地位，结果的不同是由历史路径中的特定事件决定的。当序列过程允许随机事件，每当发生这样的事件，可能的未来就会变得不一样。沿时间回溯，当前的状态是由实际历史中发生的随机和非随机事件的特定序列决定的。如果系统可以重启，结果会不一样，因为随机事件会不一样。

当随机事件与选择相结合，导致的结果则是非随机的。这似乎有些

奇怪，因为随机变化通常被认为对存在的非随机结构具有破坏性。但是当考虑各种各样的随机变化时，有一些变化可能会有利于结构的组织。如果随机变化很小，同时又有机会接受或拒绝产生的变化，并且选择是基于一致的标准，则留存下来的结构不可避免地会符合选择的标准。留存的结构会变得越来越有组织，越来越擅长做某些事情。因此，进化应当被视为在可能性空间中寻找新路径的非随机方法，虽然作为选择对象的初始变化是随机产生的。

反复的选择只有应用于指令时才能发挥出全部威力。回想一下，适当的指令能让化学和物理允许的任何结构成为可能（只要你有必需的资源）。由于反复选择能累积（从选择标准的角度看）有用的信息，因此指令会变得越来越好并且选择会不断作用于这些指令的产物，最终的产物也许会变得似乎极为不可能，像无穷小概率驱动所做到的事情一样。

指令从何而来？

如果仅从化学和物理的角度来看，一些事物几乎没有存在的可能性，理解其存在的关键在于认识到可以利用信息来说明和组织本来毫无可能性的结构。通过利用信息，能将不可能变成可能。因此我们必须知道拥有这神奇特性的有组织信息从何而来。“写1000个H”这样的指令使得本来不可能通过随机抛硬币产生的对象得以实现。第4章探讨了一些只有利用指令才能形成的结构。许多指令都是人类天才的产物，但人的大脑虽然很特别，却也不能无中生有产生目的性信息。

计算提供了一个机会，一个计算的输出可以作为另一个计算的输入。这样就有可能建立计算的链条，其中每个计算的输出都是下一个计算的输入。甚至可以让计算的输出作为其自身（相同的规则集）的输入。这个过程可以表示成一个环（图5.1）。我们称之为迭代（不断地重复）。

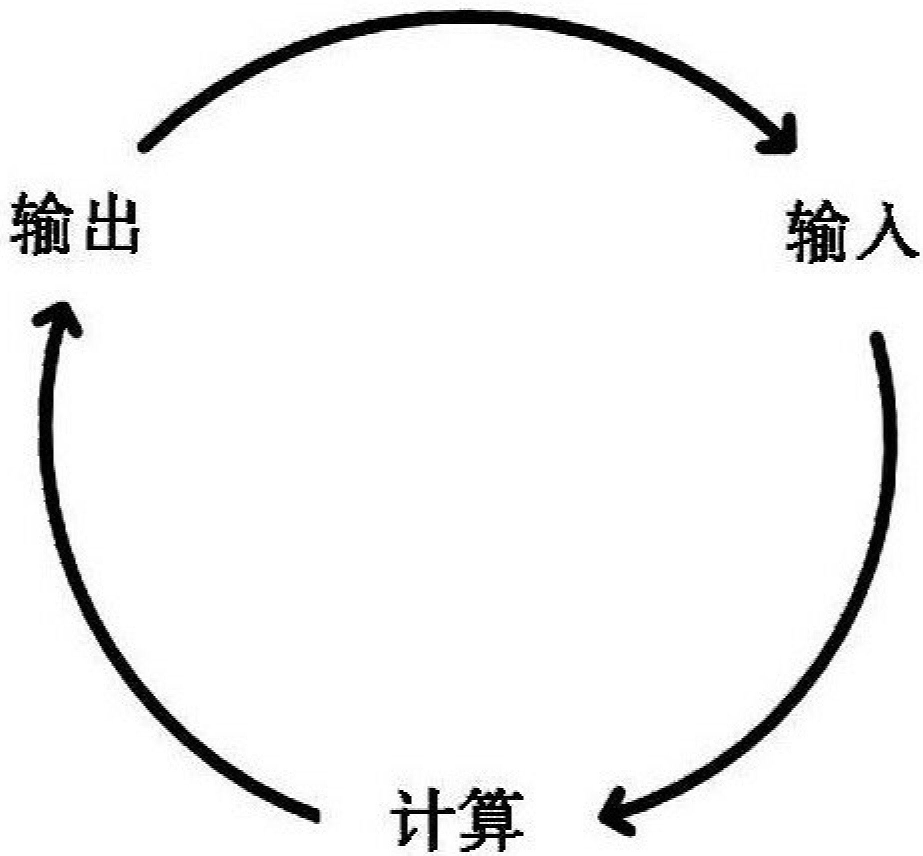


图5.1　迭代计算的概念图

大部分计算都无法迭代，因为它们的输出与输入的结构不同；但对于程序员来说设计迭代计算并不难。如果迭代计算是确定性的，会有3种可能的结局：收敛到某个最终的输出/输入不再变化；进入循环，相同的输出/输入以规则的间隔反复出现；或者形成混沌。如果计算是概率性的，则还有第4种可能：输出可能不断随机变化。

迭代复制是最基于的迭代计算。如果某个结构可以被复制，显然其复制品也可以被复制。这种输入→复制→输入→复制……形式的迭代计算说不上很有趣，但是不难做到，将复印机的输出再放到输入板上就能实现迭代复制。如果复印效果不是很好，输出会逐渐退化，变得越来越无序。无序是因为复制时产生了错误，损失了有意义的信息。迭代复制很难做到完美，因此很容易损失信息。但如果在迭代复制中加入选择，一切都会不一样。

设计迭代程序的程序员通常将输入分成不变部分和可变部分。如果规则或算法不允许改变，只有数据变化，事情会容易很多。以不完美的复印机为例，只有拷贝变化，机器本身并不变化。如果程序的规则或算法部分也随意变化，则有很高的风险产生出非法程序，导致循环停止（改变的机器不再工作）。有这种特性的系统很难用电子计算机实现，但也不是不可能。对于生命系统，程序、机器和数据相互关联，全都随时间变化。只有进化过程的逻辑保持不变，甚至执行进化的方式也会演变。

存在一个普适性的策略，允许算法和数据都随机变化，同时还能确保循环的继续。如果迭代计算中加入了程序变化的可能性，通常会有3个特点：

·程序必须实现某种复制机制，并且这种机制要允许发生错误或其他变化。

·每次循环中的错误（变化）必须很小。

·每次循环必须有足够多的输出，这样至少有一个输出相对当前循环的输入没有显著变化。

如果具备这3个特点，迭代计算就能长期进行，程序（规则和输入数据）改变但计算不会停止。由于每次循环有多个输出，所以必须有某种选择机制，因为没有哪个系统能让输入/输出的数量无限增长。我们可以将具有这些特性的通用计算策略称为“选择性迭代概率计算”（IPCS）。这种系统会表现出渐进式变化。图5.2绘制了其基本特性。内环是所有进化过程的定义性特征，也是我所说的复杂引擎。这个引擎就是IPCS，是有多重输入和输出的并行计算。

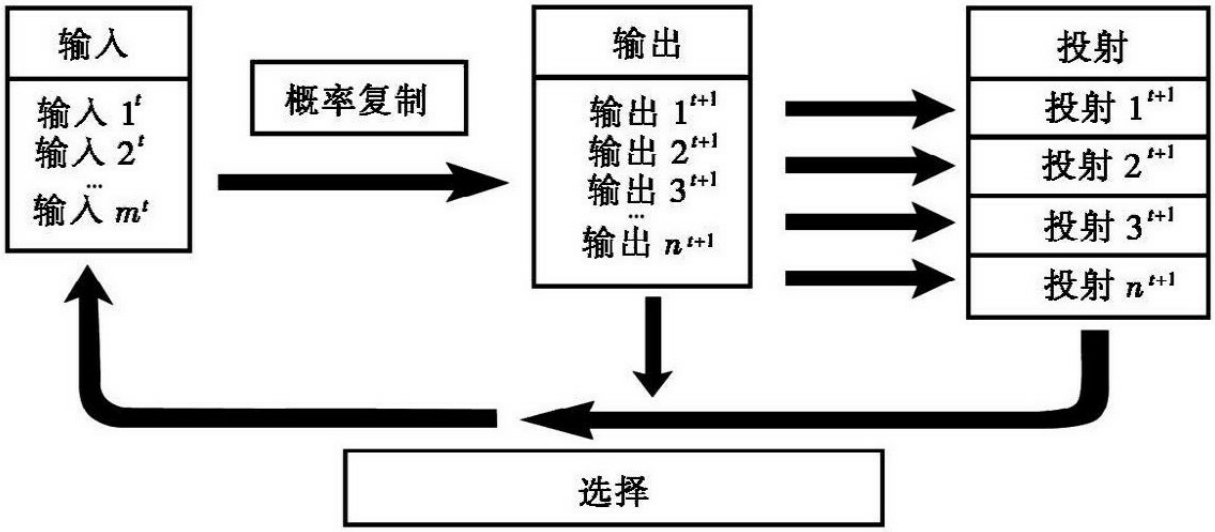


图5.2　选择性迭代概率计算（IPCS）图。这种计算策略是推动所有进化过程的引擎。上标t和t+1表示循环次数，m平均必须小于n（必须发生选择）。输入和输出都编码信息

具有复杂引擎的循环系统一般都会以输入/输出结构的形式累积信息，改进的个体被再次复制的机会更高。这个循环在随机的变化中保留更符合选择标准的变化，从而提炼信息。我称这种累积的信息具有“目的性，”因为它符合选择标准的要求。一旦出现随机变化，就会根据选择规则对产生的输出进行评估。如果变化是好的，就会成为下一轮循环的输入，如果不好，就会被淘汰。生物的选择过程混乱而随机。一些好的变化没有传递给下一代，一些不好的却留下了，但总体上个体组成的群体会加入许多更好的变化，而大多数不好的都被去掉了。IPCS计算很像一个嘈杂的机械齿轮，大部分时候都在前进。根据前面的第3条要求，至少有一个输出没有显著变化，这确保了系统在整体上不会滑坡。生命系统通常满足这个要求，突变率足够小，使得每一代的大多数个体的适应性不会比父母差太多。如果环境的辐射水平高，或者大剂量接触化学突变剂，会导致突变率过高，从而不符合要求。这样的群体会螺旋下降，每一代的适应性都不如前一代。

图5.2中标有“投射”的方框不是引擎的基本组成部分，但为了清晰起见也画了出来，因为大部分进化系统都有这个特性。在这样的系统中，选择作用于产物，而不是直接作用于输出。对于生物，输出是DNA序列，产物是生物体。对于人类活动，输出是人们遵循的指令，产物是这些指令说明的人造物或社会组织。

每次循环引入的变化必须很小，这个要求有必要反复强调。相对于输入的变化越大，得到的输出和产物的差别就越大，也意味着符合选择标准的可能性越小。大的变化导致可能性空间中的大跳跃，从而面临低概率问题。在自然界，可能性空间中“接近的”结构通常比离得远的结构要更相似；因此，与当前结构临近的结构有合理的概率得到改善。而在可能性空间中“离得远的”结构总是差别很大。而这种差别通常都是不好的，因为离得远的结构几乎肯定更加无序，无论哪种特性都会更差。为了阐明这一点，想象一架飞机。小的变化可能是将机翼的长度增加一尺。大的变化则可能是改变机翼的数量。更大的变化可能是将外形变得像一艘船或一个大金属卵石。只有小变化可能适于飞行，也只有在这其中才能找到对原设计的适应性改进。因此，找到好方案的唯一可行路径就是一小步一小步地逐渐变化。就好像在黑暗中没有扶手的舞台上之字形前进。只有步伐很小才能避免掉下去。步伐太大会有不良后果。

概率计算会引入随机变化。每当执行概率计算，就会以一定的概率产生变化。在许多系统中，这种机制就是简单的错误。生物和一些计算机程序利用更复杂的变化机制，包括对之前成功的序列进行重排、混合和匹配。无论细节如何，对于受复杂引擎驱动的计算，每一轮n个输出中（图5.2）必须至少有一个适合作为下一轮循环的输入。如果不是这样，计算就会终止。对于生物就是没有个体能够繁殖，种群灭绝。

可以非随机地选择输出作为下一轮循环的输入。本质上选择涉及产物（生物体）与环境的互动。有时候选择规则很明确，例如“死去的动物没有后代。”有时候又是间接的和概率性的，例如“可能有捕食者躲在那块岩石后面。无论怎样，具有复杂引擎的系统在非随机选择规则的作用下都会逐渐演变。在改变的过程中，系统自然而然会累积符合选择规则的信息。如果施加新的选择规则，系统又会马上开始累积符合新规则的信息。新信息的来源是被选择的随机变化。产生和保留的最简单的可能变化构成了最基本的信息单元。通过反复的选择和保存记录，系统就能逐渐累积相关的信息。如果选择是非随机的，累积的信息也会是非随机的。

需要澄清的是随机变化对于复杂引擎并不是必需的。图5.2中所展示的基本循环适用于任何变化来源。如果可能变化的范围是有限的，可能产物的谱系就是有限的，过程的创造性受限于可能的变化。随机变化提供最大的机会，如果期望的产物是已知的，指定特定的变化就能急剧

加速过程。IPCS策略的力量和美在于其能在没有预先计划的情况下表现出最大的创造性。

指令如何改变我们对事物可能性的认识？

一旦认识到指令在复杂事物形成过程中的核心作用，就能从构造事物所需的指令本身的可能性——或者说不可能性——的角度来认识事物的可能性。复杂的指令使得复杂的物理事物成为可能。通过执行指令构造出的对象的特殊性直接来自指令的特殊性。要认识到长指令有多不可能，我们必须进入大数和小数的王国。

人类对数的使用源自计数。一些原始部落的计数没有超过2 [2] 。因此我们可以推断计数的能力不是天生的。这种文化中的人无法一致地用一碗米交换一打橙子。他们能大致估计12个橙子的分量，但如果不大费周章地与另外12个橙子比对，就很容易弄错。计数让商业成为可能，会计数的商人能轻易骗过不会计数的商人。在计数发明后不久。有些人意识到也能对事物的部分进行计数。如果将苹果切成4块，对每一块苹果也能计数。

如果要计的数超出了我们的手指和脚趾的数量，一个简单的方法是给每个数起个名字。古埃及人对从1到1000之间的每个数都有专门的名字。如果要处理的事物的数量不是很大，并且也不用对数进行操作，这样的系统也能应对，但如果使用某种重复性的系统，则可以简化数的处理。例如，在语言中我们用一百二十三这样的组合短语而不是一个专门的词来描述这个数。但用组合词也很麻烦，如果将词语换成符号又可以节省很多空间。将十六万二千六百七十四写成162674显然更加便捷。更重要的是，基于符号的计数便于操作数字。十进制系统对算术非常方便。如果你认为在所有数字系统中算术都一样方便，试一下用罗马数字LIV除以XVI [3] 。算术的方便解释了为什么十进制系统现在在世界上被如此广泛地使用。

十进制系统另一个很实用的特性是能够表示我们日常不会遇到的很大或很小的数。例如，在我写下这句话的时候，地球上的人口数量估计为七十亿七千二百五十万七千三百零二，或者7072507302（而且以每秒3人的速度增长）。到中部夏令时2012年10月13日10：54为止，美国政

府的总赤字是16168864374347美元。只需适度练习，任何人都能读写这样的数字，尽管它们所表示的数远远超出了我们的实际经验。想一想，如果你在75年里每天遇到20个人，你一生大约会遇到五十万（500000）个人。如果你每秒数1美元，十小时你能数36000美元（假设你不会数丢）。一年可以数到13140000美元，一辈子大约10亿美元。数16万亿美元要花16000辈子时间。以每天20人的速度，要遇到世界上所有人，要花13000辈子时间——还要假设在计数期间没有人出生也没有人死亡。这些数字与我们可能遇到的另外一些数字比起来又微不足道。表5.1列举了一些例子。

表**5.1**　一些很大和很小的数

1升气体分子的不同状态的数量	$1\times10^{50\ 000\ 000\ 000\ 000\ 000\ 000\ 000}$
500 000个字符组成的书的可能数量	$1\times10^{800\ 000}$
宇宙中的原子	1×10^{80}
宇宙中的恒星	1×10^{22}
宇宙年龄（秒）	4×10^{16}
美国国债（美元）	1.6×10^{13}
宇宙年龄（年）	13.7×10^{9}
地球人口数	7.1×10^{9}
艾奥瓦州埃姆斯的人口数	5.2×10^{4}
扔两个骰子出现两个1的概率	2.7×10^{-2}
52张扑克出现皇家同花顺的概率	1.54×10^{-6}
被陨石砸死的概率	1×10^{-10}
抛50次硬币出现特定序列的概率	1.1×10^{-15}
普朗克长度，有物理意义的最小长度（米）	1×10^{-35}
抛1000次硬币出现特定序列的概率	1×10^{-300} 随机产
生你的DNA序列的概率	$1\times10^{-2\ 000\ 000\ 000}$ 1升
气体特定分子状态的概率	$1\times10^{-50\ 000\ 000\ 000\ 000\ 000\ 000\ 000}$

标准十进制系统也有其局限性。非常大（和非常小）的数的读写很麻烦。一般使用科学记数法来简记。根据科学记数法，1000000记为1×10^{6}（或者直接记为10^{6}），3400000记为3.4×10^{6}，0.0002记为2×10^{-4}。上标6表示小数点左边和1的右边有6个0（或6位）；上标-4表示小数点右边有4个0，然后跟着非0数。使用这种记数法很容易写出数字比如6×10^{5000}，不然就得在6后面写两页0。表5.1采用了科学记数法。

表中第1项和最后一项来自物理中的气体理论。表中第2项受这本书启发，这本书大约由500000个字符组成，使用了大约40种不同的字符（26个英文字母，空格，以及各种符号，例如/；=）。用40种符号的字母表排列500000个字符有40^{500000}种方法，也就是1×10^{800000}。这就是用

40种符号的字母表能够写出的500000个字符的书的数量。当然，其中绝大部分都是乱码；只有一小部分对懂这门语言的人有意义。这也说明了一本睿智的书有多特别[4]。

有了科学计数法，写很大和很小的数变得很容易，这是优点同时也是缺点。在科学中太大或太小的数都具有误导性。在量子力学中，无法以无限的精度同时确定一个粒子的位置和动量。如果限定一个极小的空间，在其中所发生的事情的不确定性就变得极大。因此，在很小的尺度上，所有已知的物理定律都不成立了，科学认识在这种尺度上不再成立。量子力学失效的尺度大约在$1×10^{-35}$米（普朗克长度）。传递信息最快的速度是光速（$3×10^{8}$米/秒），光通过$1×10^{-35}$米至少需要$1×10^{-44}$秒。因此从科学的角度来说，谈论短于$1×10^{-44}$秒的时间是没有意义的。小学生也可以轻松写出$1×10^{-1000}$秒，但有什么意义呢？

极大的距离和极长的时间也有类似问题。最新的天文学证据和理论模型表明，已知宇宙的年龄大约为137亿年。大爆炸标志着时间的开端，因此在$4×10^{17}$秒之前的时间是没有定义的。这之前的时间的意义是什么？同样，由于光以有限速度传播，因此天文学家无法看到发出的光到达地球的时间长于137亿年的地方。人们可以谈论更远的距离，但涉及无法探测到的事物时必须很谨慎。

再回到概率的问题，抛一枚匀称硬币得到正面的机会是1/2。如果扔一个普通的6面骰子，得到2的机会是1/6。大部分纸牌游戏都是基于概率，从充分洗匀的52张牌中发5张牌，拿到AKQJ 10同花顺的机会是1/649740（科学记数法概率为$1.539×10^{-6}$）。除非你经常玩牌，否则你在普通牌局中可能从没见过皇家同花顺。你在地上被掉下来的飞机砸死的概率是百万分之四（$4×10^{-6}$），被（小）陨石砸死的概率大致为100亿分之一（$1×10^{-10}$）。可想而知，地球人口70亿（$7×10^{9}$），已经有几百人被飞机砸死了，但只有一个可信的报告有一个人被陨石砸死。虽然存在一定的风险，但很少有人会担心晚上躺床上会被飞机或陨石砸到。这些事情发生到特定的人身上的概率太低了，无需去担心——虽然它们确实会发生[5]。

这些概率虽然很小，还是可以理解。然而，一旦考虑顺序随机事件，如果事件数量很大，出现特定序列的概率就会小到难以想象。一旦

概率小到一定程度，将其视为与0不同就会有误导性。概率小到一定程度就意味着相应事件在随机的过程中从没发生过，也永远不会发生，虽然计算出来的概率要比0大一点点。

字母序列可以简单说明这种难以想象的小概率。再来看看连续地抛硬币。每次的结果要么正面要么反面（概率各0.5）。我在桌上抛了50次，结果为
H·H·T·T·H·T·H·H·T·T·T·H·T·H·T·H·T·H·H·T·H·H·T·H·T·T·H·T·H·H·T·

由H和T组成的50个字母的序列有1.1×10^{15}（2^{50}）种可能。因此这个特定序列的概率大致为1×10^{-15}。如果地球上每个人每秒抛一

次硬币并记录结果，上面这个序列在一天里大约会再现一次。如果我记录100次而不是50次，那么全人类什么也不做只抛硬币，大约5×10^{12}年（当前宇宙年龄的400倍）才会再现一个特定的序列！抛1000次产生的序列的概率为1×10^{-300}，要抛5×10^{282}年才会再现。这样的概率与0没有区别。如果地球上所有计算机不干别的只模拟抛硬币，宇宙年龄的千万亿倍时间也无法产生某个特定的序列。换句话说，如果你写下一个1000个H和T组成的平常序列，你可以绝对肯定用随机的方法永远也无法再现这个特定的序列。这个认识揭示了概率和存在性之间一个明显但仍然很重要的关系。如果设想的对象的概率低到一定程度，它就永远无法仅仅因为随机成为物理存在。但一旦某个低概率对象存在了，我们就能生成其拷贝，而无需受制于概率问题。显然，基于随机组装假设的概率计算可能很有误导性。

现代社会不断创造和使用远长于1000个符号的序列。大部分用于制造有用事物的指令都是长符号序列；用书和光碟存储。所有这些序列都很特别。通过概率讨论我们可以绝对肯定，随机排列符号无法产生任何有意义的长指令。那长指令是如何产生的呢？说某个人创造了它们并没有解决问题，因为人如何做到的机制还是不清楚。

我们已经看到了一个可能的答案。在预先不知道所期望的序列的情况下，复杂计算引擎有能力创造很长的具有目的性的符号序列（指令就属于这种目的性序列）。事实上，复杂引擎是唯一所知的能高效做这种事情的机制。这意味着，如果我们想理解人类如何构思各种事物和创造各种技术，如果我们想真正理解复杂事物是如何产生的，我们就不能无

视这一点。

下一章将详细探讨一个特定的用计算机实现的复杂引擎是如何解决简单的“最多1”问题。这个任务是在没有将答案编码为输入的条件下创造全1的字符串。你能很容易做到这一点，因为你“知道”0和1的差别。要让计算机“知道”这个差别从而做到这一点就有点难。一个办法是编程让计算机随机生成0和1字符串，直到出现一个字符串的所有位加起来等于这个字符串的长度。这个方法的问题是要生成1千个1组成的字符串，需要生成大约10^{300} 个不同的字符串。在一个远远少于10^{100} 年只有10^{80}个原子（限制了运算这个问题的计算机的数量）的宇宙中，这件事情不可能做到。令人吃惊的是，借助于复杂引擎，我的电脑不到1分钟就能解决这个问题，而且字符串越长这个算法带来的效率提升越高！根据我们的日常经验，远远长于1000比特的指令很常见。要通过随机方法生成是不可想象的，仅有的两种可能的科学方案是：①宇宙被预先设定为会产生在过去、现在和未来要用到的所有指令；或者②它们通过一种高效的计算策略产生。在预先不知道期望的信息结构的条件下，我们唯一知道的能组装大规模目的性信息的就是复杂引擎。说某人“构想”了一长串指令只会混淆问题。“构想”不是一个机制，无法告诉我们在大脑里进行的是什么计算。

这一节我们仔细探讨了用随机的方式创造事物的可能性，从而明确了一点，复杂事物、不可能通过随机创造产生。这个原理虽然很浅显，但是很重要，部分是因为它可以用来反推。如果一个对象的随机组装是不可能的，那我们就知道它的产生不是随机的。这就是佩利的论点，但他论证的时候还不知道复杂引擎，上帝是他唯一的选项。

从哲学上来说，非随机组装意味着什么呢？一方面看，非随机地创造某物意味着先前具备的知识在其创造中起了作用。当我们写信或写文章时，我们之所以能够构造具有特定意义的特定字符序列，是因为我们对我们所写的东西已经有所认识。如果我们一无所知，我们就一个字也写不出来。黑猩猩可以被教会使用键盘，但敲击出的字符串对人类不会有任何意义。使用同样的碱基，科学家永远也无法通过在试管中合成随机的DNA序列并将其表达成细胞来创造新的生命形式，除非他们预先知道要合成什么序列。随机序列无法做到。

从另一方面看，非随机创造又意味着对可能的产物施加了约束。约

束刻画了物理世界。以栅栏为例。精心建造的围栏的立板都是垂直的，间隔均匀，顶部离地都是一样高。为什么许多围栏都是这样？用同样的板子完全可以设计出许多种类的围栏，而且大部分都很无序。之所以如此有序是因为建造者施加了约束。每块板子的钉立位置都要符合特定的标准。化学和物理定律组成的约束就如同围栏建造者的规则，只是化学和物理定律无处不在。选择也是约束，反复的选择建立了非随机性。指令是特殊类型的约束，指令越长（越复杂），能生成的产物就越复杂。因此，能够组装指令的复杂引擎为复杂事物的存在提供了一条途径，它通过一种系统的、非随机的方法组装必要的信息来实现这一点。没有这台引擎，我们所见的和所用的大部分事物都毫无机会存在。有了它，神奇的事物就变得平常。

生命是计算的产物吗？

生命的进化是被研究得最透彻和最著名的复杂引擎的实际例子。当达尔文在1859年发表《物种起源》的时候[6]，他对信息、计算、蛋白质、DNA这些现代概念一无所知。他甚至不知道基因。他的理论是基于仔细的观察和逻辑思考。作为敏锐的自然观察者，自然种群中观察到的变化以及自然的多样性给他留下了深刻印象。在所有自然种群中，产生的后代许多都无法存活到繁殖期。他问道，如果环境对一些后代的特征有哪怕一点点偏好，进化怎么会不发生呢？

后来在遗传学、生物化学、分子生物学、细胞生物学和计算机科学中的所有发现都支持进化论，从而证实了他的观察和逻辑推理。当然也不是没有反对者。少数科学家和许多非科学人士认为进化论不正确，因为它无法解释一些重要的问题。其中最重要的是生命的最终起源，以及飞行、视觉等重要创新的起源和细胞生物化学的巨大复杂性。所有研究过这些问题的科学家都承认生命的起源还没有很好的解释。达尔文自己就明确说过他的理论没有解释生命的起源，这个理论过去150年来的发展也没有从根本上改变这一点。奇怪的是为什么一个理论会因为无法解释其范围之外的问题而受到责难。这个不完备并不会削弱进化论解释在我们的星球上繁盛的许许多多生物的能力。这只说明我们还需要能用科学实验证明并且与已有的科学认识相一致的起源理论。

至于重要创新的“问题”，现在已越来越清楚这在理论上没有问题，

只是目前对动植物发育过程中器官形成的分子机制以及重要的形态创新如何通过突变和自然选择产生的认识还不清楚，对所需事件的界定又太模糊。不过这种情形正在迅速改变。在过去20年里，对器官形成的分子机制的解释取得了巨大进展。新的发现为器官如何进化出现提供了合理而详细的解释。这个领域通常被称为"EvoDevo"（发育进化学），目前是分子生物学最热门的研究领域之一[7]。

地球生命进化的最佳解释范式是DNA编码蛋白质塑造生物体（参见第4章）。所有生物都有独特的形态、生理和行为。一些特征是环境作用的结果，比如风刮断树枝或动物因营养不良瘦弱，但表型的大部分方面都是遗传自父母。

从分子层面分析，蛋白质的作用决定了生物体的表型，而蛋白质的表达又取决于DNA序列。回想一下前面的例子，T4噬菌体的保护外壳完全由蛋白质组成，因此其形状和韧性等结构特征显然也是由其成分蛋白质决定。蛋白质改变，特征也会改变。花具有特定的颜色是因为蛋白质酶在花瓣中合成色素分子；如果没有酶就不会有色素，花瓣也不会有颜色。酶改变，色素就会不一样，花的颜色也会不一样。花长出花瓣是因为调控蛋白决定了花在发育过程中细胞生长的基本"设计"图样。另一些调控蛋白则决定了花瓣形态的调控蛋白在植物体内的作用时机和位置。正是调控蛋白表达的时空变化网络最终引导了所有可遗传的动植物结构的形成。

拟南芥的DNA编码了16000种不同的蛋白质（包括副本在内总共26000种），其中大部分蛋白质都参与植物表型的多个方面。表型指的是生物体可观测的形貌特征。一种细菌能合成几千种不同的蛋白质，哺乳动物则超过100000种。即使最简单的生物体，目前科学也没有做到对每种蛋白质的特殊作用进行完整的界定，但不久就有可能出现针对特定生物所有蛋白质的表型作用的巨型数据库。

由于所有蛋白质的结构都完全由对其进行编码的基因的核苷酸序列决定，因此生物的遗传表型最终也是由DNA序列决定。DNA序列来自父母和偶然的突变。这些变化经过选择积累信息，并决定了进化的长期趋势。

重温一下生物学入门课程的内容，生命进化的7条原理：

1.生物繁殖。

2.所有生物群体繁殖的后代都多于父代。这是生物的保险策略，让生物群体能挺过艰难时期。同时也确保群体在好的时期能够扩张。

3.遗传特征由生物DNA（一些病毒是RNA）的核苷酸序列决定。

4.生物体由细胞组成，DNA是每个细胞结构的物理组成部分。因此，如果细胞存活，存储在DNA中的信息就得以延续；一旦细胞繁殖，其中的DNA就被复制；一旦细胞死亡，其DNA编码的信息就失去了。

5.个体不会进化；只有生物群体会一代代随时间进化。

6.自然群体由相同“种类”的生物组成，但群体中每个个体的DNA序列并不完全一样。这意味着不同个体生成的蛋白质也会有细微差别，产生的生物体也会有微妙的差别。遗传学家称这种差别为“群体中的遗传差异”。

7.大自然绝不仁慈。如果某个个体在错误的时间出现在错误的地方，或者无法获得足够的食物，或者找不到配偶，都很糟糕。死亡的、受伤的和不走运的没有留下后代，它们的DNA就不会延续到下一代。现在存在的所有可遗传表型都是基于成功延续下来的DNA序列。

自然界有一条绝对的规律是没有哪个种群能一直扩张。即使对于人类来说也是如此。地球是有限的。细菌的增长生动地揭示了这一点。假设某种细菌能分裂，一个变两个。10代以后，如果都不死，就有1024个细菌，增加1000倍。20代后，就会达到10亿（10^9）。如果培养基适合生长，一夜之间从一个细胞就可以培育出数以亿计的细菌，对于微生物学家来说，这很平常。但如果一直这样发展下去呢？100代后，如果没有细菌死亡，将会有10^{30}个细菌，只需132代，细菌的质量就会与地球质量相等！再过53代（总共185代），细菌的质量就会超过已知宇宙的质量。而我们只是假设细菌每代有两个后代；牡蛎能有数百万后代。人类如果不节育的话平均每对配偶会有8到10个小孩。如果人类从现在起每一代的数量增加一倍（平均每对配偶4个小孩），人类只需1175年（47代）就会超过地球的质量！显然这是不可能的。要么我们自己有意识地控制后代数量，要么大自然替我们这样做。

从中得出的一条自然规律是，生物要么限制繁殖，要么大部分后代还没有繁殖就会死亡。所有例外都只是临时性的。由于很少有生物会有意识地限制其后代数量，与环境的互动会决定哪些后代能存活到繁殖。在自然界这个选择有两种方式，要么纯粹是运气，要么是因为一些生物更善于生存和寻找配偶。生物通过无穷多种途径与环境互动，互动的细节决定了它们是否会挨饿，或被吃掉，以及是否能繁殖。由于自然种群中的差别，必然会有一些个体更擅长某些事情，一些则做得更糟。一旦这些事情影响到繁殖，就会影响未来群体中呈现的基因。

这个故事的背后是蛋白质和基因。基因决定蛋白质，蛋白质又决定表型；因此如果一个可遗传特征会给后代带来优势或劣势，相应基因在种群中的比例就会增加或减少。如果基因在种群中的比例增加，就意味着下一代会有更多个体携带这种基因，并表达出相应的蛋白质。即使是很温和的选择，力量也很惊人。假设某个基因序列能带来1%的繁殖优势或劣势，很小，如果不进行细致的统计研究谁也不会注意到。群体遗传学为计算这样的DNA序列的命运提供了方法。多代之后的结果取决于种群数量和其他一些参数，如果我们假设群体数量为10000，配偶选择随机，那么具有1%优势的DNA序列会在大约2000代后遍布整个种群。如果是繁殖迅速的细菌，则只需要1个月，如果是橡树或人类，则需要40000年。同样，具有1%劣势的序列，如果开始分布很广泛，在差不多的时间里也会消失，如果初始比例很小，则会消失得更快。100万年足以发生许多变化。

只要自然种群会产生遗传性变异，一些特征会不可避免地增加，一些则会逐渐消失。由于优势或劣势主要由与环境的互动决定，即便是很小的环境变化也会导致一些特征的优势地位改变。在很长时期内如果环境的变动不太剧烈，对特征的选择会导致整个种群的基因成分漂移；如果某种特征比如亮色皮肤、奔跑更快或更好的抗病性受偏爱，则这些特征最终会遍布整个种群。

图5.3将生命进化描绘成了一个循环过程。选择通过成功的繁殖实现，繁殖的成功又取决于个体与环境的复杂互动。很显然图5.3中的循环就是图5.2中的循环的特例。由于DNA编码信息，因此图5.3也描绘了计算。DNA编码的信息既是循环过程的输入也是输出。细胞机器将这个程序转化为蛋白质，然后蛋白质进入细胞或生物结构执行计算。生物的一个本质功能就是繁殖。DNA复制过程中的错误会在每次循环中引入

（通常）很小的改变。平均孕育的后代的数量会多于父代数量。由于遗传性变异对于进化的发生很关键，除了简单的错误，还演化出了各种生物机制以确保有足够的新变异和限制产生过大的变化。这其中包括染色体的随机配对和重组等重新排列DNA序列的机制，这些将在后面讨论。

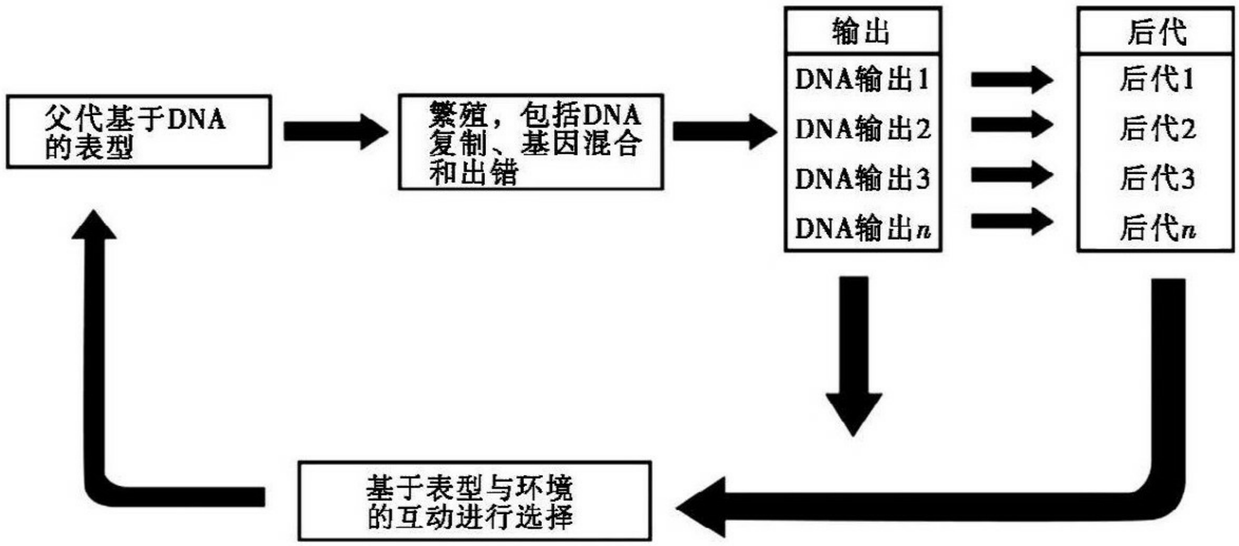


图5.3　生物进化的计算循环

因此，如果聚焦于DNA循环，很显然生命的进化就是计算；从而所有生命都是计算的产物。参与这个计算的DNA序列点滴累积核苷酸信息。累积的信息产生细胞和身体结构，以及（对生物）有益的行为倾向。累积的信息的产物与环境互动。成功用繁殖的成功来衡量。后面我们还会看到，计算机程序中也可以加入类似的性质。

图5.3描绘的是达尔文进化论。但也可以看出图5.3就是图5.2中的计算策略的具体例子。因此也可以说图5.2描绘的是广义进化论。复杂引擎比标准生物进化论更为广义，因为它不限于生命。它对任何能执行计算的系统都起作用。这也意味着图5.2中描绘的复杂计算引擎也可以用来定义非生命个体组成的群体的进化。就如我们将看到的，在判断计算机程序、社会或思想是否会发生真正的达尔文意义的进化时，这个定义会很有用。

第6章　算法进化

计算机能学习吗？

在手机上打字很麻烦。为了让打字方便，很多手机都能猜你要打什么字。打得越多，就猜得越准。这是通过学习你的习惯实现的。计算机学家已经发明了很多方法可以让计算机学习。其中许多都是基于复杂引擎的变体。在生命的世界里，进化源自对环境信息的累积——也就是学习。当DNA序列产生微小的随机变化，自然选择会留下那些能产生更“适应”环境的生物的DNA。进化出适应性表明系统获得了其适应对象的信息。鱼适合在水中生活，鸟能飞翔，这都是因为与其生存环境相关的许多信息被存储在它们各自的DNA中。这一章关注的一个方面就是这种适应信息的来源。我们将看到实际上只有两种可能的来源：信息要么是来自其适应的对象，要么就是随机产生，然后选择那些能提高适应性的变化来提取信息。无论哪种情况都可以将这种获得信息的过程视为学习。

计算机与人类大脑相比有两个优势：很快，而且能准确地复制信息。但学习不仅仅是复制。学习涉及抽象。当你下载一个程序或一段音乐，你的计算机会存储数据的准确拷贝。计算机很擅长做这种事情。

但是当你学一首歌或上一堂课时，发生的事情很不一样：你在创建表示这首歌或这节课的神经连接模式，这种模式能引发特定的行为，并存储新的记忆。你在头脑中创建的神经模式与你的老师或作曲者头脑中的模式并不完全一样。我们擅长学习抽象的思想，而不擅长准确的记忆。如果你不相信这一点，可以去访问一下某个事故或大事件的目击者。

计算机学习的概念很难界定，但容易认识到。它必然涉及获取新的信息，并且不仅仅是简单的记忆。学习一词意味着获得推理的能力。学到的东西让计算机可以解决以前从未遇到过的问题。这是智能的标志之一。因此，计算机学习也被称为人工智能。

一个常见的学习任务是模式识别，比如在人群中识别一张脸。机器在学习识别模式时有一个简单方式是人工神经网络（ANN）。在计算机科学中，神经网络一词指的是一种以有序方式组织的逻辑系统，如图6.1所示，并不是指的基于生物神经元的系统。

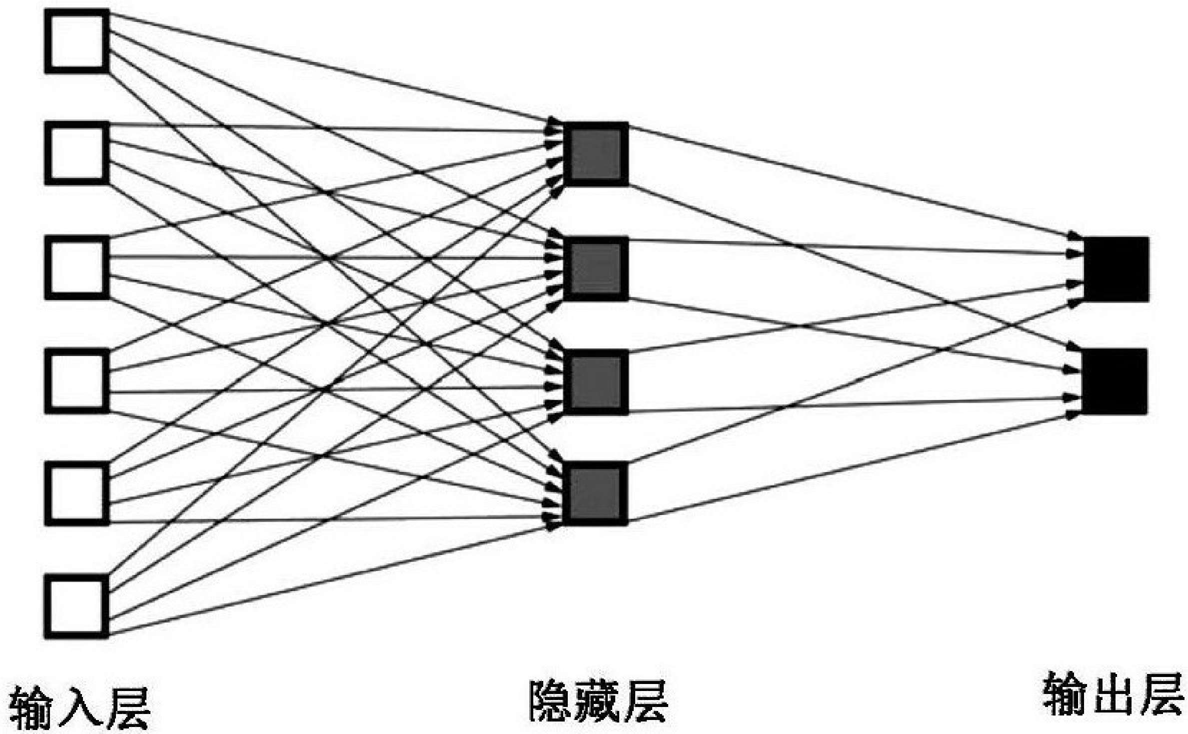


图6.1　人工神经网络（ANN）。箭头表示节点之间具有方向和权重的连接。方块表示节点

神经网络由3个特征刻画：节点、连接和权重。节点（图中的方块）是接受输入，计算，然后通过连接输出的逻辑装置。节点相互连接，每个连接都有权重，权重决定连接对下一节点的作用效果；权重越大，作用就越大。图6.1展示的网络有3层，上一层的每个节点都连接到下一层的每个节点。一个ANN至少要有2层，也可以多于3层。上一层的节点不一定都要连接下一层的所有节点，有时候还可以有反向连接。图6.1中的ANN只有前向连接。

ANN的运作原理如下。电压或符号0和1形式的信号（信息）输入到输入层。这些输入一起组成了需要处理的数据集。每个输入节点的信号通过图中用箭头表示的连接继续传送到隐藏层的节点。信号在通过连接传送时会乘以权重因子。权重可正可负；若权重为零，连接就不传递信

号。隐藏层节点根据某个简单规则对输入进行加权得到输出信号。例如，可以取输入的均值。处理后，隐藏层节点会把信号送到输出层。这些信号也会乘以权重，然后输出层根据这些加权输入计算一个值。最后的输出组成“答案”。

这种方式的输出是确定性的；它们完全取决于输入值和权重。一些ANN能产生复杂的输出（很多输出节点），一些则只有简单输出（一两个输出节点）。图6.1中的ANN就是两个输出节点。输出可以是两个数或两个电压，可以解读成是或否，或者是、可能和否。例如，如果输入描述的是一幅数字图像的像素，那么是或否的输出可能回答的是如下问题：图中是你的姨妈吗？

要让ANN能正确识别出图片里的姨妈，就必须进行训练。训练的目的是寻找合适的权重，这样一旦输入具有某种特征，就能产生特定的输出模式。一般是给ANN各种输入，一些是正面例子，一些不是。如果目标是识别出你姨妈的照片，权重就必须调整到可以对有姨妈的图像输出是，对没有姨妈的图像输出否。让人吃惊的是，有可能找到这样的权重组合（只要所有有姨妈的图像具有某种独特性）。通过训练，ANN可以从没有见过的图像中识别出有姨妈的图像。现在有很多商业软件已经应用了这种学习方式。

寻找合适的权重有两种不同的策略。一种原理很简单，只需对权重进行随机的微小变动，然后使用表现更好的作为下一轮随机变动的种子。也就是说进化出合适的权重。这种策略很有效，但是很慢。还有一种方法是利用所谓的反向传播计算权重。反向传播需要对实际输出和期望输出之间的差别进行量化。还要计算节点输出对各种输入权重的敏感度。如果能够对各节点进行这种分析，就能在每次训练后计算新权重，让新的输出更接近期望值。训练越多，表现就越好。

与对随机变化进行选择相比，反向传播更快，因此在条件允许时通常被采用。但如果条件不允许，例如不知道目标输出值，或者没有最优答案时，基于随机变化的策略就是唯一的选择。有一个例子是用ANN帮助设计更好的机翼。不知道最好的答案，但在引入新的设计约束时，适当训练的神经网络能给出供测试用的解决方案。

ANN只是用于机器学习的许多策略之一[1]。学习的目标基本都是

将一个大的可能性空间分成两个或多个子空间，让所有的正例都属于其中一个子空间，所有的反例都属于其他子空间。对于ANN来说，空间由所有可能的权重和输入组成。就算是规模不大的ANN，可能的权重和输入的组合的数量也很大，无法一一验证。这时必须发展抽象规则。所有机器学习策略都涉及反复的训练，大部分都会在训练中调整参数，计算出比之前的值更好的新值。如果理想值可以直接计算出来，就不用训练了，也不需要学习了。答案直接计算出来。

机器学习算法让计算机可以从训练集获取信息。要做到这一点最简单的办法就是记住训练集中的所有例子，但这需要存储很多数据，而且一旦遇到的问题没有包含在训练集中就没办法了。要解决没有遇到过的问题，必须对信息进行概括后再存储。受过训练的神经网络的权重就是概括信息的例子，通过适当训练，神经网络就可以用这些信息正确分析没有遇到过的输入。

反向传播是一个很有用的确定性技术，但只有可以用数学方法将外部数据的信息融合进权重时才有可能。许多问题无法构造成这种形式。如果不知道或者无法从期望的答案中提取所需的信息，那还有什么信息来源呢？答案出人意料：随机猜测加选择。随机猜测能为任何问题提供答案，问题是要找到一个有效的方法来利用这个缺乏组织的信息源。

如果学习策略完全依靠随机猜测，在大多数时候，好的答案会淹没在糟糕答案的海洋中。显然要解决这个困难得想办法进化出答案。事实上许多实际问题都有多个答案，一些答案比另一些要好一些，而且好的答案可以通过微小的变化进行改进。投资策略、机翼和运输网络只是其中一小部分例子。工程师们遇到的许多问题都有这个特点。

细微的改进有一个很重要的特点是它们通常具有一定的可能性。复杂引擎就是利用了细微改进具有合理的可能性这个优势，虽然大的改善基本不可能。通过这种方式可以很容易找到好的答案，我们可以通过一些简单的进化算法的例子来认识这一点。

什么是进化算法？

不是所有进化都与生物有关。20世纪50年代，计算机学家就创造了基于生物学原理的算法，但是与细胞中发现的分子结构并没有关联。到

20世纪70和80年代，密歇根大学的约翰·霍兰和他的学生对这类算法进行了深入研究[2]。其中一些程序的表现非常类似生物进化，虽然它们只不过是计算机软件。这类算法的存在证明了进化过程并不仅限于生命系统。它们也揭示了进化、学习和计算的密切关联。

进化算法的细节各不相同，但都遵循以下6条原理：

1.它们都包含可能对某件事有用的信息（信息体）的多个拷贝。通常这些信息表示的是对某个问题的试探性解答。2.这些信息体是“一代一代的”。每一代都有源自前一代的新拷贝。

3.信息体从一代到下一代之间会随机变化。结果使得群体成员并不完全一样。

4.所有变化的范围和幅度都很小。不受约束的大的随机变化没有用。

5.在每一代中，一些信息体复制得更加频繁。

6.必须有一致的标准来根据各信息体所编码的信息决定哪些信息体可以复制以及有多频繁。

算法是遵循特定顺序的逻辑步骤。进化算法有一个特点：它们包含图6.2所示的循环过程。

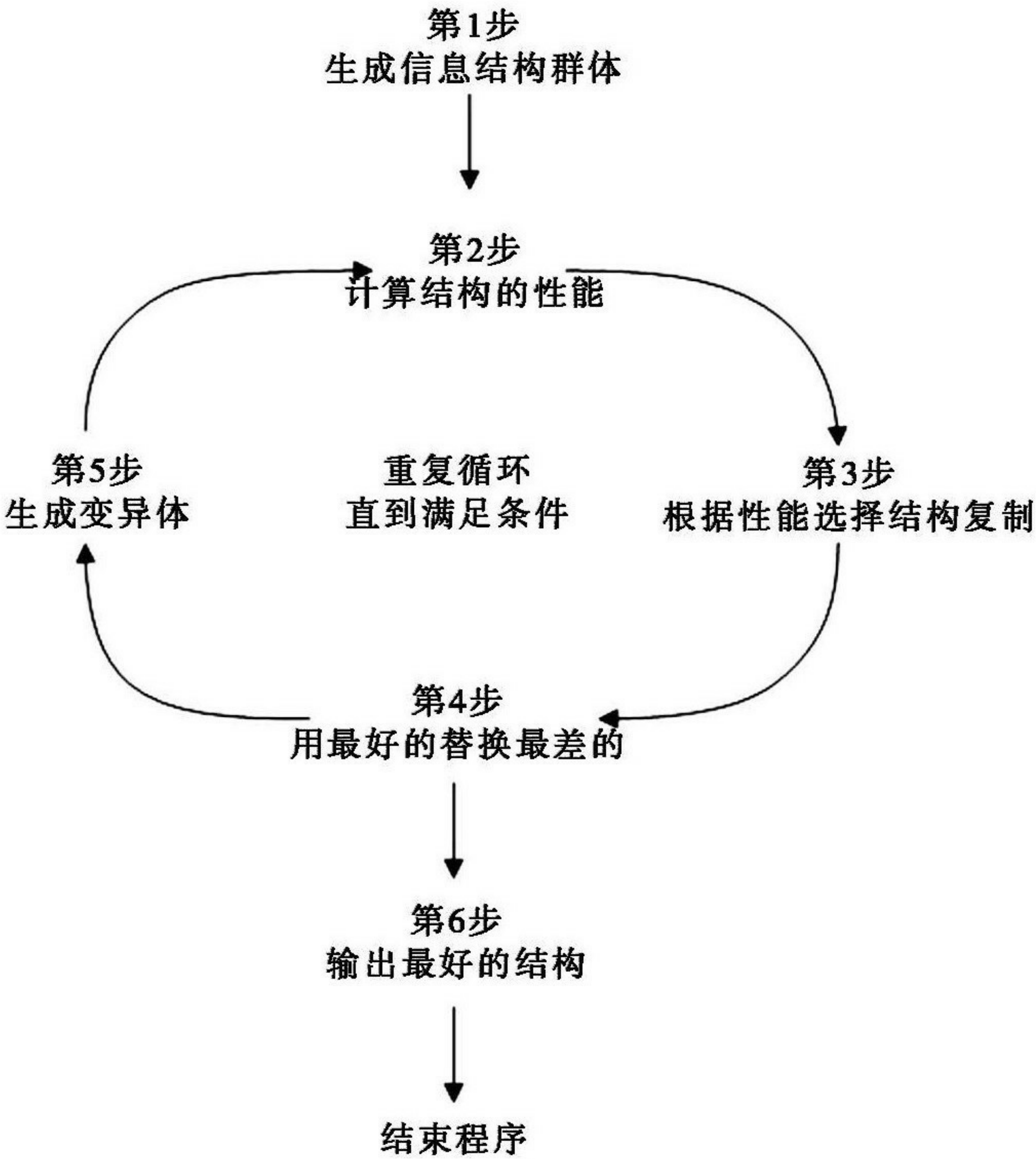


图6.2　进化算法的基本结构（根据Ashlock 2006修改）

算法的第1步是生成初始信息体结构，通常是随机生成；可以将其视为试探性答案，不需要很好。第2步是评估。根据某个标准对所有信息结构进行评估。评估是必须的，而且非随机。第2到第5步反复进行，从随机变化中捕捉有用的信息并累积。第5步是新信息的来源。由于只

有最差的结构被替换，就算所有突变拷贝都比父代差，系统整体也不会“滑坡”。总的效应是累积能提高成绩的随机引入的信息。如果在某一轮循环中没有出现改善，也没有关系；总会有一些循环改善。（根据选择标准）反复基于每一代最好的进行改进来实现长远的改变。一旦达到预先设定的目标或时间，程序就输出最好的结果然后停止。

只要所评估的对象发生变异，这个通用方案就能为许多问题产生出“性能”越来越强（评估成绩越来越好）的结构。事实上任何编码信息的体系都能作为进化算法的基础。计算机科学家给出了许多这样的例子，这一章探讨其中两个。同样，对性能的评估以及选择复制的信息体的规则也很多。一致性是关键，并且第5步生成的变异体不能与父代差别太大。如果新的变异体与前代差别太大，就很难带来改善。另一方面，如果变化太小，也什么都不会发生——每一代基本都与前代一样。

有适量的变异，每一代就有合理的可能产生出一些性能强于前代的信息体。一旦出现，有更好评估成绩的信息就能传递给后代。通过反复执行这个策略，整个信息体群体针对目标问题就会变得越来越好。

图6.2给出的方案通过突变复制并留下最好的信息体确保每一代都不会差于前代。这个特性并不是必须的。在许多应用中，如果突变率不是太高，群体中所有成员都可以突变，算法仍然能正常工作。要产生出程序员自己也没有想到的变异，变异至少在一些方面必须是随机的。

进化算法的一个子类称为遗传算法，因为它们细致模仿了生物进化。信息被编码为符号的线性序列；并且群体中的所有成员都有低突变率。突变会随机改变单个符号（例如，随机从1变成0或从0变成1），并且通常会加入交换机制，在信息体群体的不同成员之间交换信息子集（类似于生物中的有性繁殖和基因重组）。还有一些生物学机制也会被用到，比如复制或删除序列中的一段。

遗传算法有一个共性就是突变率有“最佳击球点”。如果变化机制太保守，变化就会很慢，如果突变太快，系统就会掉进随机的混乱，不能保留前代的记忆。如果突变率被调节到合适的值，群体就能迅速进化，每个成员携带的信息都能越来越符合评估标准的要求。下面我们通过两个例子来了解遗传算法的工作原理。

例1：最多1，一个很简单的问题

这个计算机程序将1和0组成的任意序列转换成全1序列——对于人来说很简单，但对于不知道如何识别“1”的计算机来说并不容易。这个问题清晰地揭示了遗传算法的工作原理。图6.3描绘了一个遗传算法，可以在没有明确告知计算机如何“识别”0和1的情况下完成这个任务。为了简单起见这个例子中的序列只有10位，但完全可以是任何长度。

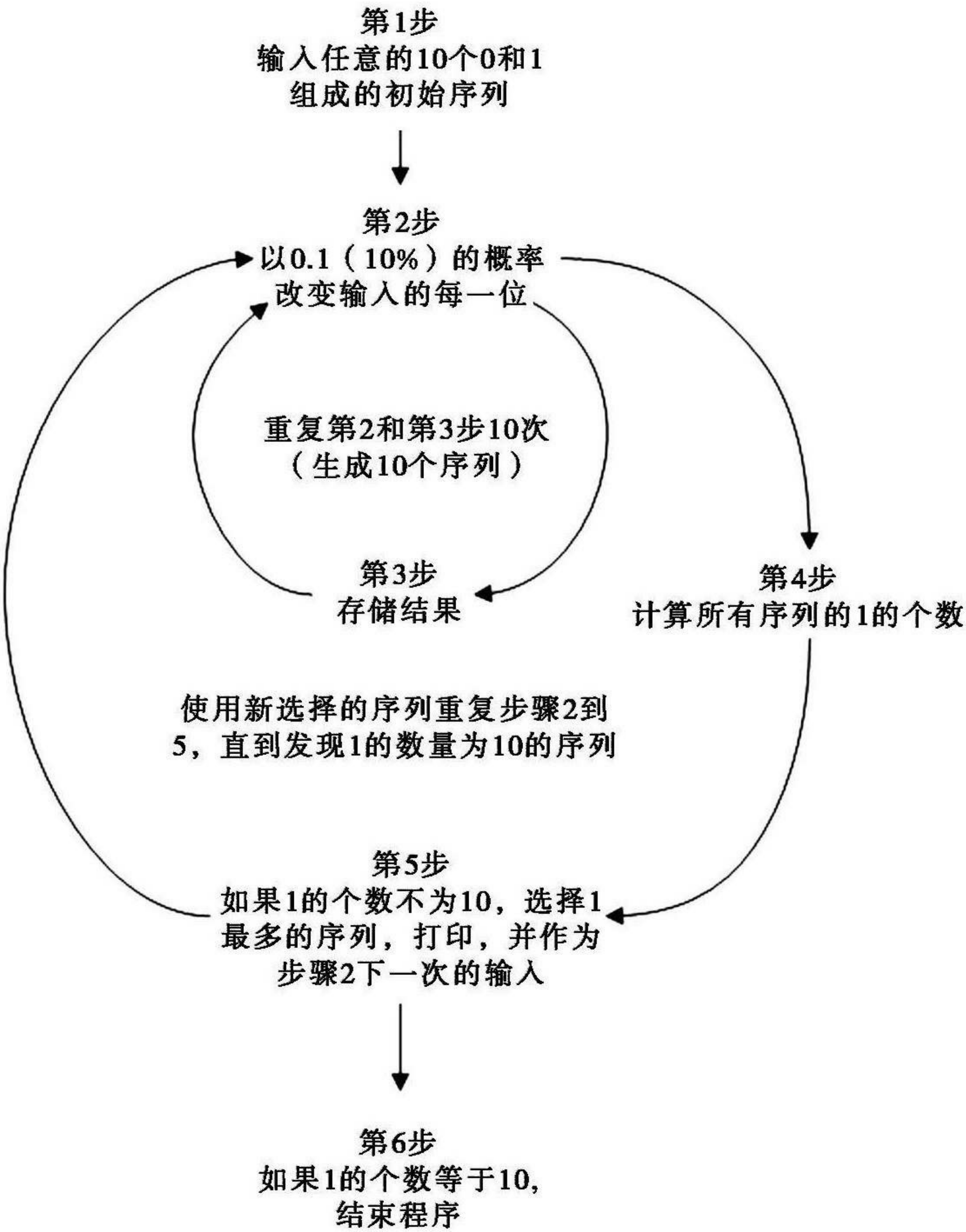


图6.3　解决最多1问题的遗传算法

在这个算法中重复2、4和5步骤的循环就是图6.2中进化算法特有的循环。重复步骤2和3的循环为每一代生成10个尝试结构。举个具体的例子：假设算法的初始输入是1001000100（1024种可能的01序列中的一个）。表6.1记录了从这个序列开始的一次计算的过程。最终的输出就是所预期的1111111111。

对于表6.1中给出的每一轮计算（迭代），在复制输入字母生成输出时，有0、1或2位被随机选中翻转（0变成1或1变成0）。也就是说每次复制字符时有10%的可能产生变化。为了方便起见，表中列出了每个变体的1的计数。每一轮结束后将计数值最高的字串（粗体）选出来作为下一轮迭代的输出。例如，迭代2那一列所有字串都是从迭代1中计数值为5的字串产生的。这次实验用了5个回合（迭代）输出10个1的字串。其他实验解决这个问题的回合数不一定相同。每次实验的细节都会有所不同，因为变化（突变）是随机的，但最终的输出总是由10个1组成的序列。1的“来源”是将一些0变成1的随机翻转。这个过程也会随机地将1变成0，但这些变化在选择过程中被舍弃了。

表**6.1**　基于图**6.3**中的算法从序列**1001000100**开始的计算过程

迭代 1		迭代 2		迭代 3		迭代 4		迭代 5	
字串	计数	字串	计数	字串	计数	字串	计数	字串	计数
1001100100	4	1011010101	6	1011011101	7	1011011101	7	1101011111	8
0001000100	2	0001010101	4	1111011101	8	0111011101	7	1111011011	8
1001000100	3	1001110100	5	1011010101	6	1111011101	8	1011011111	8
1001100100	4	1001010101	5	1011111001	7	1111011111	9	0111010111	7
1001010101	5	1001000101	4	0011011101	6	1101001101	6	1111011110	8
1011000100	4	1001010101	5	1011101101	7	1110011101	7	0111011101	7
1001000101	4	1000010101	4	1011011111	8	1110111101	8	1111111111	10

迭代 1		迭代 2		迭代 3		迭代 4		迭代 5	
1001010110	5	1001010111	6	1011011100	6	1111010101	7	1101011111	8
1001000100	3	1101010101	6	1010011101	6	0111011001	6	1111011111	9
1100000100	3	1011011101	7	1001011001	5	1111011101	8	1110111111	9

表 6.2　　比较用遗传算法或随机方式生成全 1 字串所需的时间（必须生成的字串的数量）

字串的位数	遗传算法需要生成的字串数量（平均）	随机方法需要生成的字串数量	进化算法的加速比
10	60	500	8
50	600	10^{15}	1 666 666 666 667（$1.7 \cdot 10^{12}$）
100	1330	10^{30}	10^{27}
1000	34000	10300	$3 \cdot 10^{297}$

表6.1中的运行过程总共生成了50个01字串。反复试验表明平均约为60个。如果不是对现有的字串进行细微的改动，而是每次都生成全新的字串，则大约需要50个回合（总共500个字串）才能产生出全为1的10符号字串。因此，搜索10个1组成的字串时，进化算法比简单的随机搜索大约快10倍。对于更长的字串，节省的时间会更多；表6.2体现了这一点。表中第1行就是这里讨论的例子。

1千比特没有多少信息，然而如果完全依赖随机，要生成编码这么多信息的特定字串，所需的时间远远大于已知宇宙中原子数量乘以以纳秒（10亿分之一秒）计的宇宙年龄。这意味着如果宇宙中的原子每纳秒生成1比特，在宇宙存在的时间内都不可能产生所期望的字串。因此我们可以确定不可能随机生成1000个1组成的字串。然而用最多1算法却很容易。这展现了复杂引擎的魔力。

要把效率提高到如此高的程度，突变率的设置很重要，必须依据字串长度进行优化。当突变率接近1/n时（n为输入字串的长度）通常能表

现出最优性能。图6.4描绘了字串长度为100比特时突变率的影响。

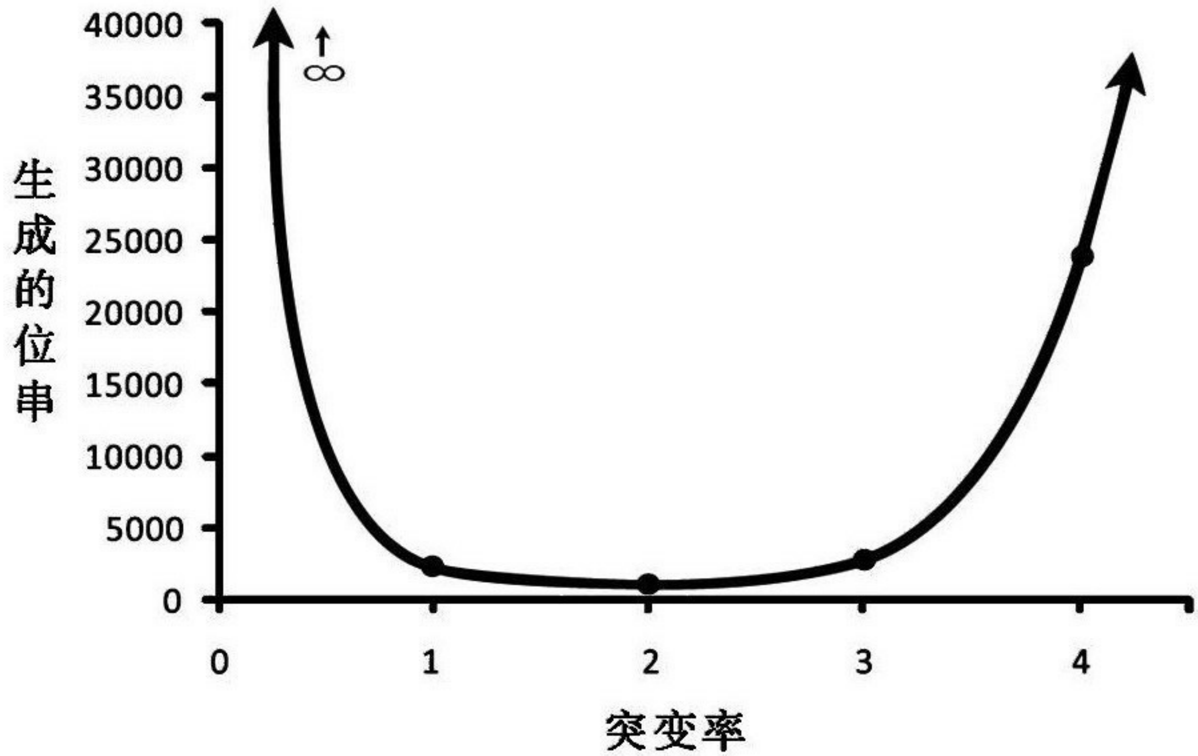


图6.4　最多1遗传算法（图6.3）找到100个1组成的位串需要生成的位串数量与每一代平均改变的位数（突变率）的关系。突变率为0时算法一直运行，突变率100%时大约需要生成10^{30}个位串。突变率为2%时只需要生成大约1　300个位串

最大1问题很简单，人们不会想要用进化算法来解决它，但这个问题很好地揭示了进化算法的工作原理，很适合用来入门。

例2：塔尔塔罗斯

丹·阿什洛克和他的学生研究了能优化格子板上虚拟推土机行为的程序。这个游戏名为塔尔塔罗斯，最初由阿斯拓洛·特勒于1993年提出，当时他是卡内基·梅隆大学的研究生[3]。图6.5中给出了一个例子。推土机的任务是将箱子推到格子板的边和角上。推土机没有传感器，一次只能推一个箱子。推土机一次走一步，每步有3个动作可选。推土机可以前进一格（F）、左转（L）或右转（R）。命令组成的序列，例如FFRFL，就是决定推土机行为的控制序列。这个序列指令推土机依次前进两格（FF），然后右转（R），然后前进一格（F），然后左转

（L）。在塔尔塔罗斯中推土机和箱子不能同时占据一个格子，两个箱子也不能在同一个格子。如果推土机往前进入一个有箱子的格子，箱子就沿推土机前进方向被“推移”一格，除非前面被另一个箱子或墙挡住。如果动作要求推土机推多个箱子或墙，什么也不会发生，算法会继续执行控制序列的下一个命令。

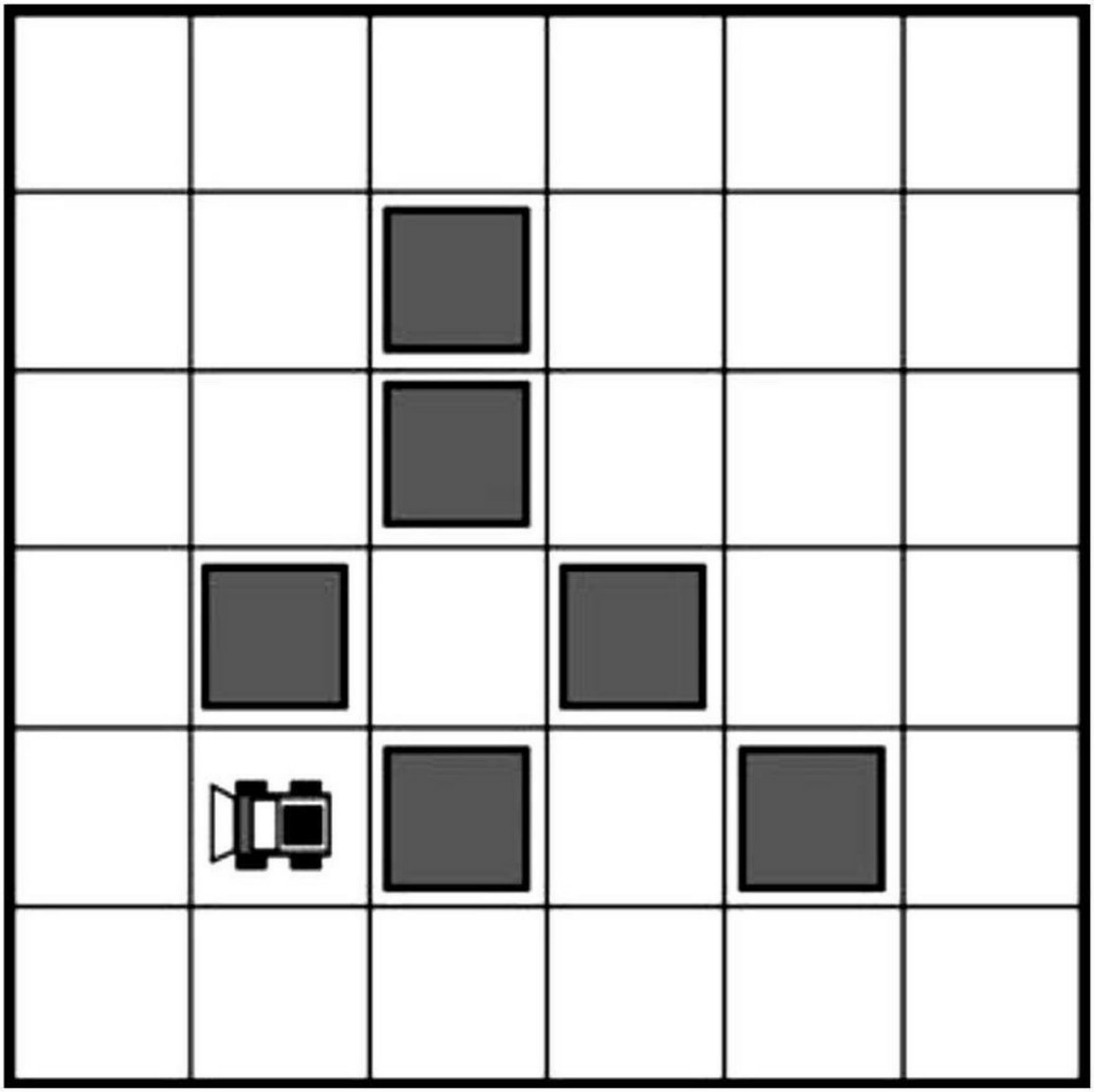


图6.5　6×6的塔尔塔罗斯板，有6个箱子和1个推土机。这个图是根据丹·阿什洛克提供的图修改的（参见注释3）

遗传算法能产生出用有限步获得高分的控制序列（6×6的板子一般80步）。每推一个箱子到墙边得1分，墙角的箱子得2分，中间的箱子得0分。图6.6给出了6×6的板子和6个箱子的4次不同结果。这个游戏的板子至少要3×3，并且箱子数量远远少于格子数量。也可以有多台推土机。

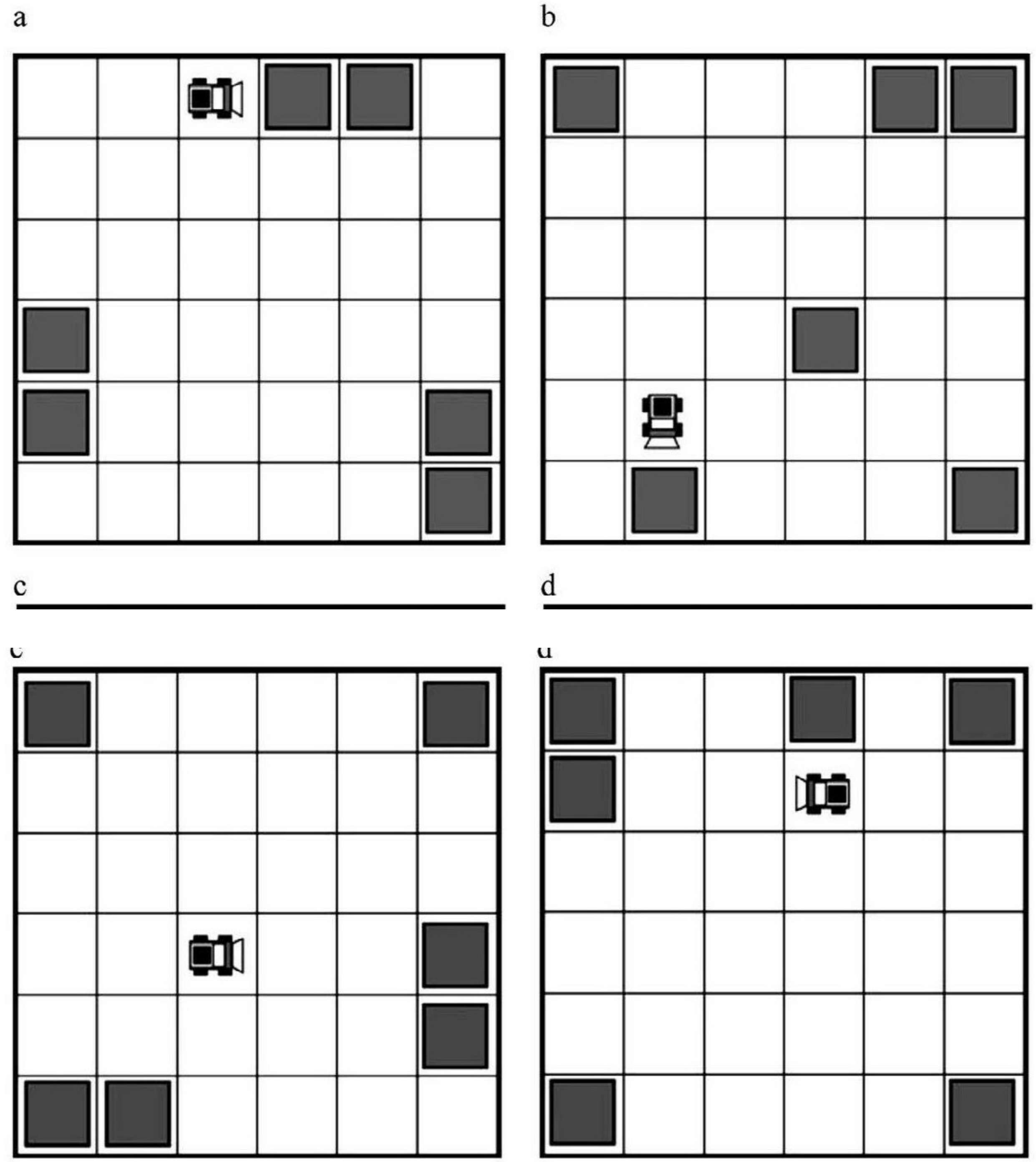


图6.6　4个80步之后的6×6的板子。a板：7分；b板：8分；c板：9分；d板：10分。根据丹·阿什洛克提供的图修改（参见注释3）

这个游戏有2个不同版本。在简单版本的塔尔塔罗斯中，要为给定的初始状态（例如图6.5）进化出最佳答案。在第2个版本中，推土机的行为要能够对随机的初始状态都能表现得很好。无论哪个版本，都是生

成多个控制字串，用板子进行测试，选出最好的进行突变然后再次测试。

如果使用固定的初始状态，突变后运行一次就能够对控制字串的表现进行评估；算法如果运行足够长时间总能找到最优的控制字串。如果用随机生成的板子测试控制字串，则需要生成多个板子进行一系列测试，然后计算控制字串的平均得分。完成测试后，选出平均表现最好的控制字串作为基础生成新的字串。突变可以是简单改变控制字串的一个字符，也可以包含其他变化策略，比如选择性复制和/或交叉（也称为重组）。无论哪种情形，都需要维持一个控制字串群体，并且将表现好的控制字串作为后代字串的“父母”。应用这个算法会逐渐发现越来越好的控制字串（分数越来越高）。对于所有的遗传算法，新信息和变革的来源都是对个体的（微小）随机变化进行选择。

如果板子的状态随机生成，游戏会很难，如果给推土机增加记忆和传感器可以显著改善性能。结果表明，推土机记住之前的动作会很有用，例如是否推动了箱子，或者推了无法移动的箱子，或者推到了墙。增加记忆和传感器会让算法更加复杂，但分数也会更高。

规则能够进化吗？

在前面的两个例子中，程序的输入（01串或控制字串）在每次循环中会产生微小的随机变化，但程序本身并不变化。而进化算法的一个子类，遗传程序，则是让程序可以像输入一样进化。这样就可以像改善输入/输出一样改善算法。

完成一个任务所需的步骤可以用许多方式表示。遗传程序常用一种名为分析树的信息结构。分析树用树形结构表示操作之间的关系。图6.7给出了一个简单任务的解法的3种不同表示。任务是计算a和b两个变量的乘积然后加3。图中给出的每种表示可以用于不同的场合。公式很紧凑直观。算法是计算机可以执行的命令序列，而分析树则揭示了过程的逻辑结构。

公式：

y=（a×b）+3

算法：

第1步输入变量a和b的值

第2步将a与b的值相乘

第3步将乘积加3

第4步打印y=“第3步得到的值”

第5步结束程序

分析树：

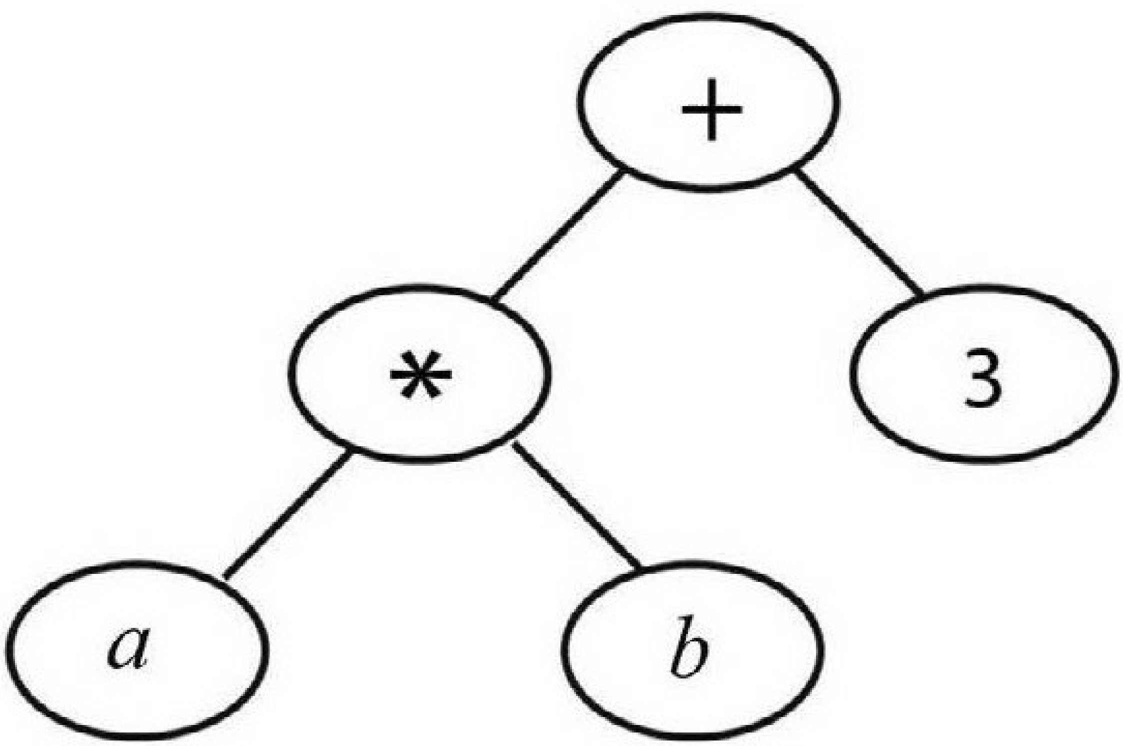


图6.7　用公式、算法和分析树表示同样的逻辑操作

分析树有两种节点，终端节点（或叶节点）和操作节点。图中a、b和3是终端节点，*（乘）和+（加）则是操作节点。终端节点下面不连接任何节点。节点分层，并且最多与一个上层节点和两个下层节点相连。分析树的大小就是节点数量。有且仅有一个节点具有特殊性质，它

没有上层节点，并且直接或间接与其他所有节点相连。这个特殊节点称为根节点。图6.7中根节点是+节点。

分析树自底向上解读。图6.7中的树包含3层，最底层包括终端节点a和b。它们连接到第2层的操作节点*。这个结构的意思是a与b相乘，结果放在*节点。第2层包括两个节点，*节点（现在包含的是a乘b的值）和终端节点3。它们连接到顶层的操作节点+。意思是3与*节点保存的值相加，结果放在+节点。由于没有比+更高的节点，+就是根节点，操作完成（答案放在+节点）。

所有公式和算法都可以表示为分析树，并且分析树可以通过维持一个分析树群体进行进化，随机改变节点和连接，然后选择分析树作为新分析树的“父母”。随机改变分析树的一个有效方法是选择节点然后删除节点下面连接的整个子树，将被删除的子树用随机创建的新子树替换。另一种方法是交换群体中两个成员的子树。图6.8展示了（生物）交换的分析树版本。

即使是相对简单的算法，如图6.3中那种，分析树也有很多节点，不方便绘制。但如果问题类型合适，并且巧妙应用，遗传程序就能进化出各种分析树，对庞大的算法空间进行探索，为一些极为困难的问题找到创造性答案[4]。

很显然包括遗传程序在内的进化算法都是复杂引擎（图5.2）的具体实现。它们都涉及群体、突变和选择。在后面的几章中我们会看到一系列其他实现。但在此之前，我们先了解一下如何将进化过程可视化。

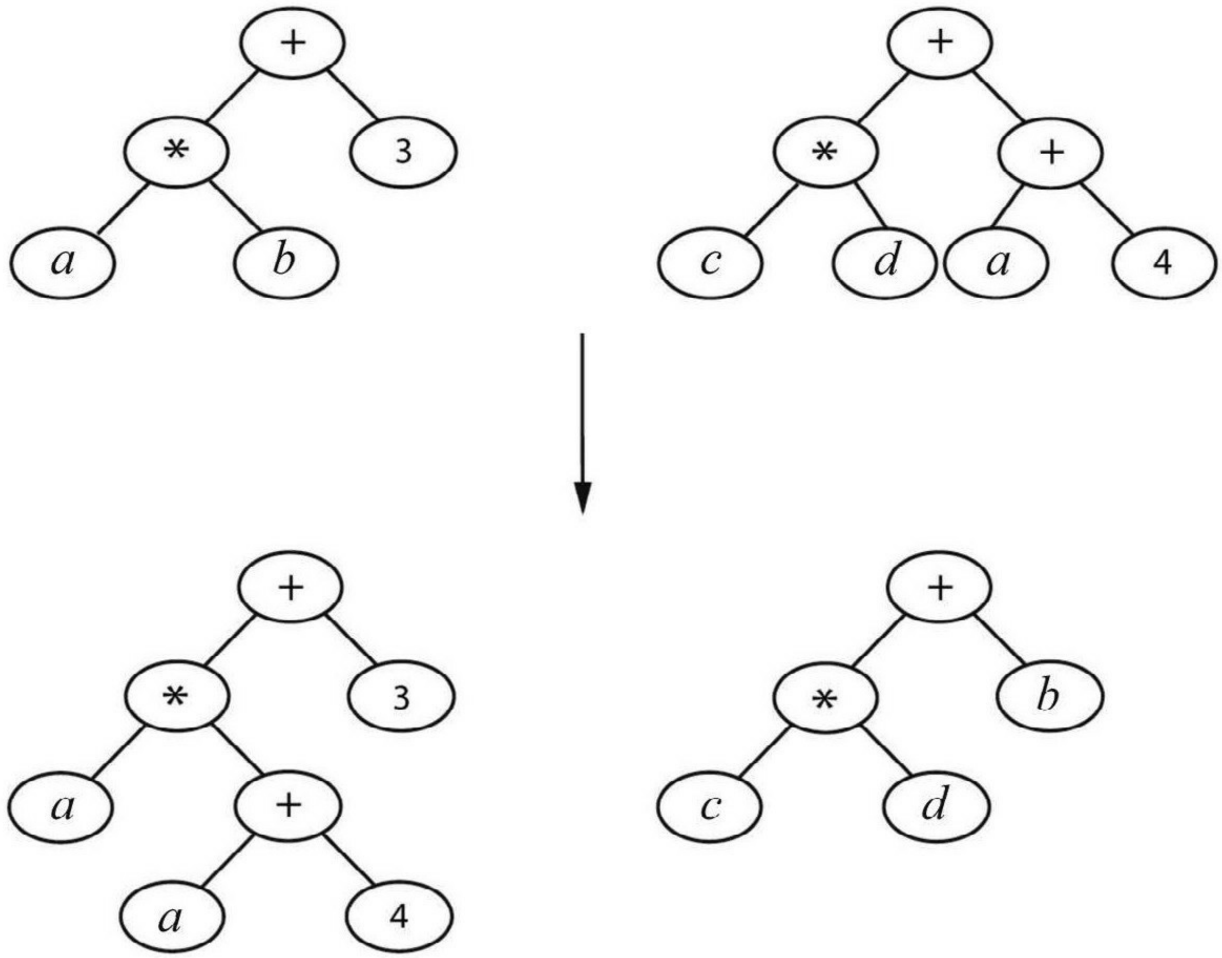


图6.8　分析树的交换。右上角树的子树“a+4”与左上角树的子树“b”交换，产生出下面的两棵新树

如何将进化计算的过程可视化？

进化算法和遗传程序本质上是不断寻找和优化问题的答案。为了看清这个过程中发生了什么，想象由某个问题的所有可能解答组成的一个抽象多维空间，每个可能的解答都是这个空间中的一个点。这种空间通常都大得难以想象，但是可以通过所谓的适应性地形来大致地可视化，适应性地形用3维地貌来显示问题的大量可能答案的值或某个性质。1932年休厄尔·赖特首次提出用这种图形帮助阐释生物系统的进化 [5] 。它对进化算法尤为适用。图6.9给出了一个简单的数学函数的3维图，这个函数因为显而易见的原因有时候被称为“山峰函数”。这幅图通过显示与x和y值对应的z值，让我们可以看见方程的解空间。

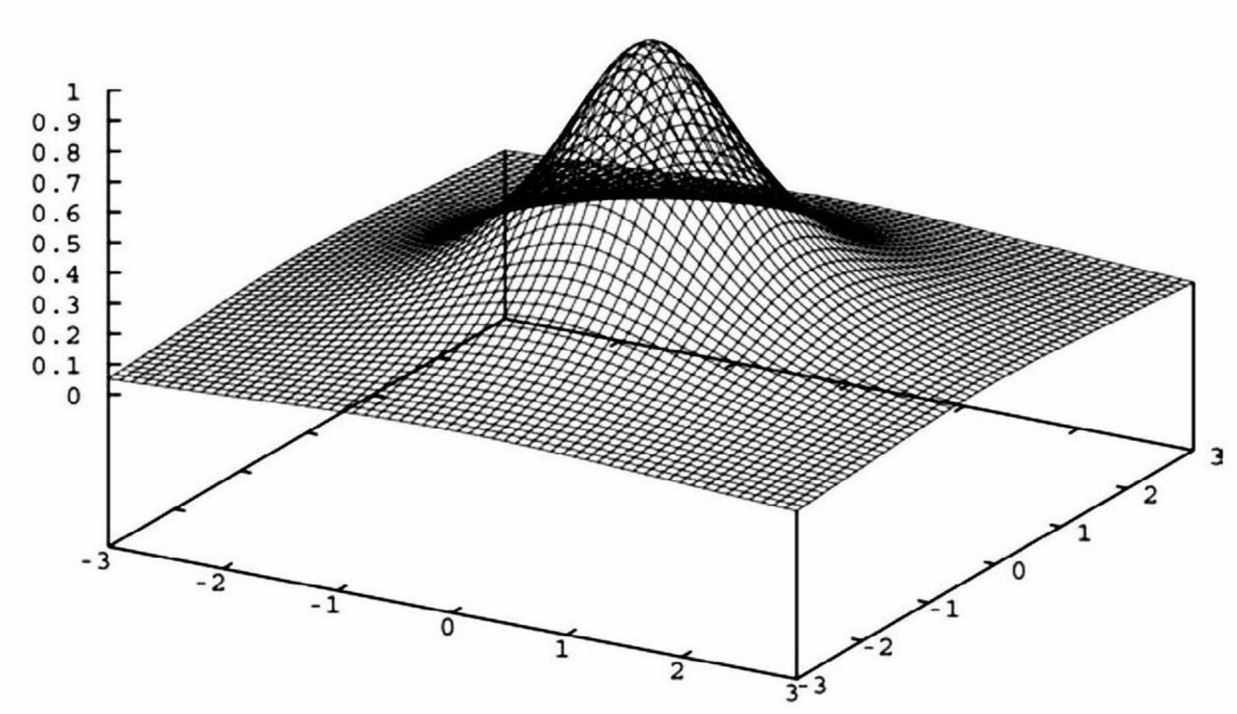


图6.9 z=1/（x^2 +y^2 +1），“山峰函数”的3维图像。感谢丹·阿什洛克提供图片

无论代入的x和y值是多少，z的最大可能值是1。这可以用微积分或进化算法证明。进化算法并不知道解微积分问题，但有把握找到这个方程的最大值在哪里。这个算法创建一个猜测的“群体”，分别进行评估，选择最好的（最大的值），然后在最好的猜测“附近”创建一个可能答案的新群体。我们用算法的一次运行输出来说明这个过程，这个算法维持20个猜测值组成的群体。每个猜测值都是一对x和y值。为了简明只给出了每一轮生成的20个猜测值中的5个，并且只保留了两位有效数字（小数点后第1个非0值的后两位）。表6.3给出了运行过程。只显示了第1、第3、第4、第5和第7回合。每轮最好的猜测（最高的z值）用粗体显示。初始猜测介于1000和-1000之间，随后的猜测限制在4倍于最好猜测大小的区间内。例如，如果某回合y的最好猜测是3，则下一回合猜测的y在-3到+9之间随机选取。z值越大，答案越好。可以注意到每回合的z值都比前一回合更接近1.0。这次运行到第8回合时（没有显示），z的值到了0.998，舍入为1.0。如果不舍入，要达到正确的1，需要100多回合。

再看图6.9，可能的答案就好像散布在区域地形上，最好的答案位于山顶。最初的猜测没有哪个特别靠近山顶，但有一个比其他的更好。下一回合就在第一轮的最佳答案附近进行猜测。这回又会有一个比其他

猜测更接近理想答案。只用了5个回合，就找到了“答案”0.94，到第7回合，就找到了0.990。这个算法要很久才能找到正确答案，但一定能找到。

表**6.3**　搜索山峰函数最大值过程中进化算法生成的一些值

		猜测 1	猜测 2	猜测 3	猜测 4	猜测 5
回合 1	X	180	100	940	-60	680
	Y	60	-20	-320	740	-700
	Z	0.000028	0.000096	0.0000010	0.0000018	0.0000050
回合 3	X	17	-21	59	6.7	9.6
	Y	-42	-8.4	25	8.4	-34
	Z	0.0048	0.0019	0.00024	0.0086	0.00079
回合 4	X	12	3.0	-3.7	13	-0.49
	Y	17	2.5	7.2	-3.0	15
	Z	0.0022	0.061	0.015	0.0052	0.0042
回合 5	X	8.8	0.72	-1.9	3.4	0.24
	Y	4.6	-2.2	5.0	-1.5	0.10
	Z	0.010	0.16	0.034	0.069	0.94
		猜测 1	猜测 2	猜测 3	猜测 4	猜测
回合 7	5 X	0.60	1.0	-0.069	0.24	-0.038
	Y	0.070	0.37	-0.049	-0.060	0.19
	Z	0.98	0.94	0.990	0.94	0.96

无论是哪种进化算法，都要在（解空间中）目前已知的最佳猜测的附近进行猜测；如果不这样就没有什么机会获得稳定的改善。如果不限制在最佳猜测的临近区域，算法就变成了纯粹的随机猜测，而大部分问题都基本不可能通过随机猜测找到好的答案，因为这些问题的解空间都很大。对于山峰函数，需要上百万次随机猜测才能找到最佳答案，而进

化算法只需不到200回合，而且从上面的例子可以看到，算法只用了5个回合就找到了一个相当不错的答案（0.94），虽然找到最佳答案的回合数要多得多。

山峰函数与很多问题都有一个共同的特点，大部分可能的答案都离峰顶很远，但在任意特定答案的邻域中，几乎有一半答案比当前答案更接近峰顶。因此，只要将猜测限制在上次猜测的邻域，就几乎总是有新的猜测会有所改进，搜索也就可以推进到越来越好的答案。而如果不对搜索进行约束，绝大部分猜测都不会比当前的猜测好。只有当可能的答案数量或搜索空间足够小，全局搜索才有可能成功。要让进化算法能顺利工作，每个回合或多个回合中至少要有一个答案在山坡上爬得更高。通过不断往更好的答案推进，才能最终发现峰顶的最佳答案。

山峰函数很容易用遗传算法解决，因为其形状很光滑。许多问题都没有这么简单。解空间经常会有许多峰和谷。图6.10给出了一个假想的例子。这种情况下进化算法很容易陷入一个低矮山峰，永远也找不到最高峰。前面说过，几乎所有解空间都是多维的。图6.9和6.10显示的是2维解空间（x-y平面）。10位的最大1问题的字串解有10维解空间。1000比特位的串则需搜索1000维空间。高于3维的空间没法画出来，因此图6.10只是对高维解空间的一个有用但并不完美的想象。

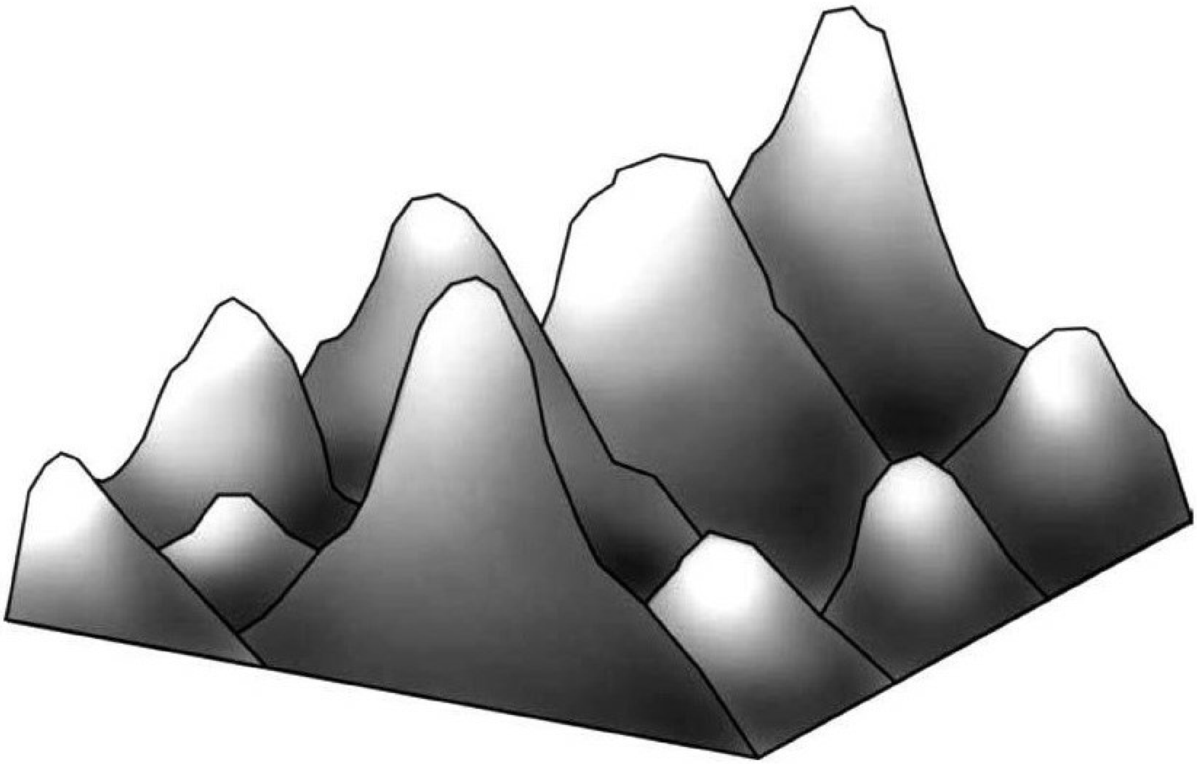


图6.10　由许多峰谷组成的2维适应地形。每个峰都是一个局部最优的答案，谷则代表了糟糕的答案。对于进化算法来说找到一个山峰很容易，而要找到最高峰则通常很难

随机对于学习是必须的吗？

这一章有4个要点。一是进化并不仅限于生物领域。二是解空间（状态空间）通常都很大，纯粹随机地搜索答案注定会失败。三是往往可以进化出随机猜测发现不了的好答案。四是如果需要新的信息又不知道到哪里去找，随机会是一个很有用的来源，而且经常是唯一的来源，但也需要对搜索管理得当。

从计算的观点看，对事物的操作可以是确定性的也可以不确定，或者两者兼而有之。确定性过程的结果可以从开始获知，而不确定过程的结果只有到实际发生之后才能知道细节。不确定一词意味着随机性对结果起作用。就算99%肯定的结果也是不确定的，因为没有100%的肯定。

各种背景下使用的进化算法，以及更广义的机器学习，面对的解空间都具有一定的相关性，好的答案聚在一起，最佳答案附近的答案没那么好但也不错。因此学习算法可以将多维解空间划分为多个区域，将好

答案与差答案区分开来，从而允许进化算法成功“登山”并发现越来越好的答案。

如果你对答案一无所知，随机猜测将是唯一的选择。在大多数情况下我们对答案有一些了解，根据已知的信息进行适度的随机猜测往往能找到一些不错的答案；如果我们继续在“附近”（解空间）随机寻找答案，就能找到更好的答案。重复这个过程通常能很快找到相当不错的答案。如果你想要最好的答案，可能要很长时间，但经常还是能找到。

通过有约束的随机猜测（在附近的可能性空间取样），找到更好答案的机会大得出奇。只要跟着越来越好的答案，就能很快找到相当不错的答案。这就是复杂引擎的“秘诀。”很多生物进化的批评者认为，一些事物不可能进化出来，因为概率太低了，这是因为他们没有明白这一点。所有学习算法都利用了自然界潜在的有序性。复杂引擎也不例外。数学家能构造出不那么有序的解空间，从而让进化算法不再有效，幸运的是我们的宇宙不是这样构造的。

归根结底，所有成功的学习算法都要获取信息，复杂引擎也不例外。学习算法并不一定必须借助随机性，但如果没有随机输入，学习过程就是确定性的；学习的内容要么内建在算法中，要么从其他来源引进。在解空间中受到适当约束的随机采样与选择结合是获取信息的有力途径。就复杂引擎来说，可以认为这个算法的功能就是从随机性中提取有用信息。

第7章　身体内的进化

单个细胞如何形成人体？

来到光之城巴黎的游客大部分都会去卢浮宫。在卢浮宫的希腊和罗马文物区，绕过著名的断臂维纳斯，在达鲁阶梯的上面，有一尊美丽的雕像，名为萨莫色雷斯的胜利女神。雕塑虽然没有了头和手臂，仍然很迷人。雕塑家是怎么创造出来的？据说伟大的艺术家能想象作品最后的样子。大脑中的图像可以作为蓝图，但蓝图又是如何转化成雕塑的呢？具体的过程因材料而异。黏土雕塑有很多种做法，石头雕塑则只能切削。无论哪种，最终的雕塑都是添加或移除材料的繁琐过程的产物。每一次添加或移除都让形状与最终目标更为接近。每一个中间状态又成为下一次添加或移除的基础。

自然以类似的方式塑造事物，但没有蓝图。雪花的形成过程中，水分子倾向于附着到与空气接触最多的部位。每一个水分子的附着都会改变随后的分子的附着目标。基于少数简单的规则，构造出复杂的外形。岩石山峰的形成类似石雕。大地抬升，雨水和冰冲蚀掉一部分地貌，留下尖峰和山谷。自然的雕塑没有预先设计，但形成的过程与3维艺术的创造有类似之处。

雕刻的技术含量不高。原始人会用木头和石头雕刻小物件。古希腊人是这种简单技术的大师，创造了胜利女神、断臂维纳斯等不计其数的美丽雕塑。他们没有计算机可用于雕塑的3维设计，或是直接用工业机器人激光烧蚀大理石块来生成他们的设计。胜利女神可能有草图，但不可能像制造螺丝刀那样有详细的指令，或是今天用CAD/CAM（计算机辅助设计/制造）程序生成的设计图。第5章介绍了复杂引擎：收集说明目的性结构的信息的计算过程。雕塑家也会用同样的策略吗？我认为在雕塑的背后的确有类似的东西，但不完全是复杂引擎。与图5.2中的循环相比，缺少的要素是多次反复复制。

累积结构的形成在自然界随处可见。大多数自然事物都是通过对已存在结构的反复添加和删减产生。这个逻辑与复杂引擎的差别不大，但

是缺少多重复制。前面简要介绍过计算链条的思想。其一般形式为：信息体1→操作→信息体2→操作→信息体3，等等。复制在这个链条中没有扮演重要角色；父代信息体为后代信息体提供了直接基础。雪花形成时，“操作”是将水分子添加到生长的结构上。信息内在于所有物理结构中，因为信息和结构就是同一事物的一体两面。因此，从信息的角度看，所有物理过程都是计算链条。如果链条具备了反复的循环，就具有了复杂引擎的许多要素，但不一定有多重复制。我们姑且将缺乏复制的累计循环称为“构造引擎”。胜利女神和雪花都是构造引擎的现实例子。

石雕的过程中也有选择。雕刻家在每次抡锤之前，都会考虑凿子放置位置的多种可能。对下一个点的选择需要比较当前结构与想象的最终结构。水分子附着到生长的雪花上也有选择——每一次附着都有许多水分子趋近生长的雪花——但这里的选择不是基于与想象的形状或蓝图的比较，遵循的是统计性的简单规则。

为了从总体上把握事物形成的机理，我们将事物的各种形成方式组织成图7.1所示计算过程的层次。除了宇宙的起源可能是例外，所有结构的形成都是从某种预先存在的结构开始，因此顶层基本包括了一切事物。第一级分为确定性过程和概率性过程。概率性过程含有某种随机活动。与分子热运动有关的事物都划分到这部分。确定性过程不允许任何随机活动。这一类过程包括大部分电子计算和许多机械设备的操作，例如发条钟表。随机部分又分为序列性和迭代性。迭代性过程是反复执行同样的行为，而序列性过程则是顺序执行不同的行为。物理世界的大部分过程都是序列性的，有一些例如晶体的形成则是迭代性的。迭代分支又分为涉及复制的现象和不涉及复制的现象。各分支的最底层分为基于简单规则的结构和需要复杂规则（例如指令）的结构。简单规则和复杂规则之间没有明确的界线。确定性过程也可以分为序列性的和迭代性的，但这种区分不会带来新的认识，因此被略去。

复杂引擎位于图中右下角；也就是包含反复复制的概率性现象所在的位置。构造引擎所在的分支是不包含复制的概率性迭代。雪花和大理石雕像都位于这里。

构造引擎能生成很复杂的结构。用不包含复制的反复选择构造事物的最神奇的例子也许是从受精卵发育人体。身体的形成过程很类似雕刻，但没有雕刻家，也有点像雪花的形成。人体极为复杂。成年人大约

由10万亿（10^{13}）个细胞组成，包括200多种不同的细胞类型和不计其数的微小变化。细胞形成组织、器官和器官系统，它们在总体上表现出的目的性细节比受精卵DNA中能直接编码的要多得多。每个人都是从单细胞开始，DNA编码的新生婴儿的信息都存储在那个细胞中。因此科学家和哲学家们想知道，既然没有足够的DNA可以编码详细的指令告诉每个细胞变成什么样子和到哪里去，一个简单的细胞又是如何发育成人体的。

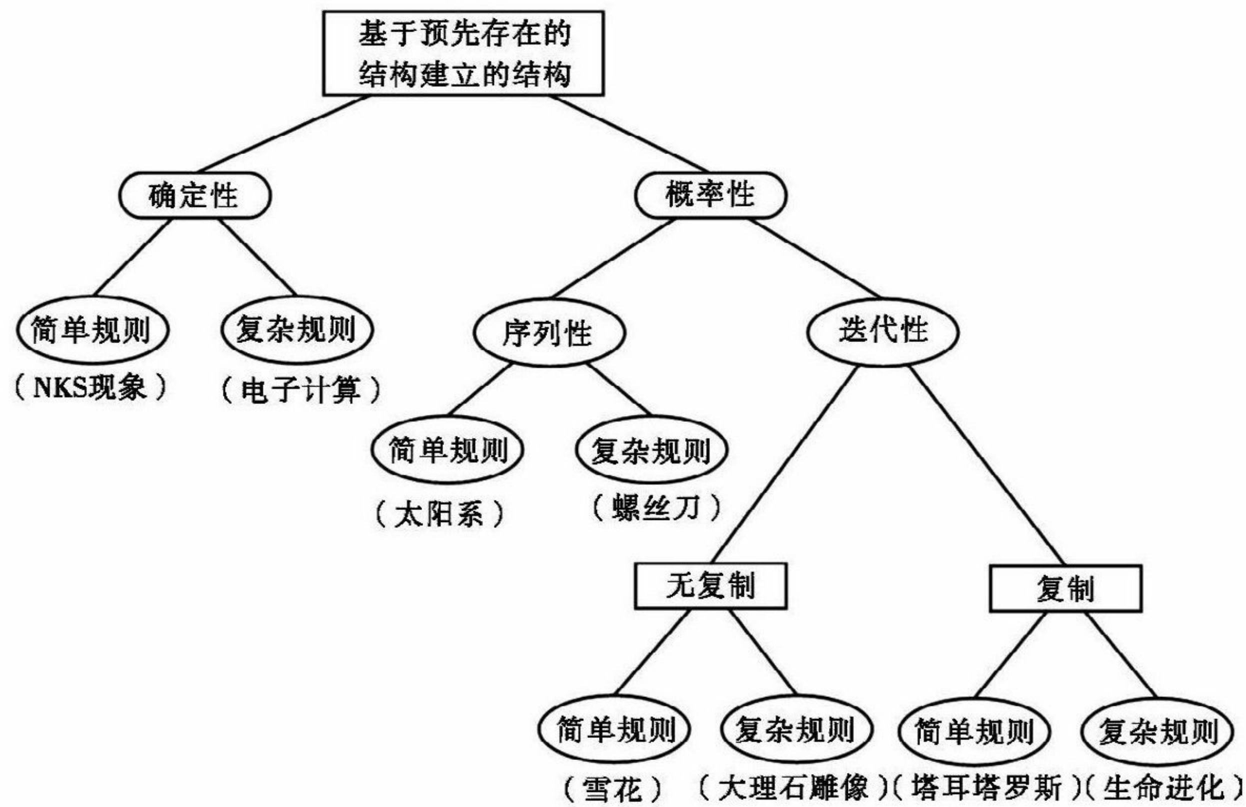


图7.1　结构的层次。各末端分支下面的括号中给出了例子

对此我们已经有了大致的认识。科学家发现，身体中的细胞根据环境信号分裂、改变蛋白质表达模式和死亡。大部分时候环境就是其他细胞。DNA提供了蓝图，或者应当称之为草图，结构则是通过添加和剔除细胞形成，逐渐构造出最终的形态。大致上，身体的形成是大量繁殖细胞然后剔除不符合局部标准的细胞的过程。在这个过程中，所有细胞都受制于一些简单的局部规则，细胞的历史和位置决定了细胞的命运。DNA则承担了规则库的作用。

以人的手为例，成年人的手由大约1000亿（10^{11}）个细胞组成，占身体的1%。这些细胞组成了皮肤、肌肉、骨骼、韧带、神经、血管、毛囊和指甲根（生长指甲的组织）。每种组织都包括几种不同的细胞类型，并且构成特定的形状。手有5根手指，每根都由骨骼支撑，通过灵活的关节连接。手的动作依靠肌肉收缩。神经连接到脊椎，由大脑协调肌肉的运动。所有结构都被皮肤包围，所有部分都配有血管。由于人的手都很相似，并且明显不同于狗的爪子或鱼的鳍，因此其设计是由人类的DNA决定的。

手比螺丝刀显然要复杂得多，我们在第4章看到，螺丝刀的形成需要大量指令。人类DNA还编码了腿、肝脏、大脑和耳朵等许多身体部位的信息，因此人类的DNA只有一小部分可以专门用于编码形成手所需的信息。显然身体中各处的肌肉、骨骼、血管和皮肤组织基本都一样，可以用同样的基因说明。但即便如此，还是必须有大量信息用于说明手与脚或鼻子的区别。这些信息具体是如何使用的还不完全清楚，但有一些是明确的。首先，手不是直接形成最终的形态，而是从大致相同的细胞中逐步产生出来。

形成手所需的大部分信息说明的都是位置怎样决定细胞的命运。邻近的细胞决定了胚胎细胞的行为。有两个过程扮演了关键角色。细胞分化，随着各种基因的开和关，细胞分裂并改变性质；以及细胞的协同运动，将细胞群放到新的环境中。除了最初的阶段，动物的发育主要是细胞的迁移，让细胞与新的邻居接触。这种迁移很随意，细胞经常“迷路”。迷路的细胞通常有两种结局：要么最终找到正确的位置，要么死去[1]。发育过程中许多细胞会死去，因为它们所处的时间和位置从周围的细胞接收到的分子信号与它们预先编程的需求不符。这些细胞可以说是自杀。

这种策略看似很浪费，但从保存信息的角度来看却很高效。要么身体要能容忍许多细胞位于错误的地方，要么就要有一个数据库告诉每一个细胞在什么时候出现在身体中的什么位置。如果有这个数据库，原则上迷路的细胞能知道什么时候从哪里到哪里去，或将自己重构成适合新位置的细胞类型。据我们所知，这种信息没有编码在DNA中，如果编码所有这些信息需要很多DNA。

基于位置的细胞选择有一个典型的例子就是手指的形成。在发育的

早期，胚胎手的形状就像桨。在这个桨中，出现了交替的基因活动模式，导致形成9个交替的细胞区域，具有不同的基因表达模式。第一个区域的模式刺激细胞的分裂和分化，导致结缔组织的形成，邻近区域的模式则导致其中的细胞自杀。交替相邻区域的细胞有的成为骨骼，有的死亡。5个区域形成骨骼，间隔的4个区域的细胞则死亡。结果是在桨的位置形成了5个分开的部分（胚胎手指）。然后皮肤细胞迁移到形成骨骼的区域周围，形成分开的手指。从信息的角度来说，“剪开”和“戴上手套”是产生这种复杂形状的有效方法。

在动物的发育中，选择性的细胞死亡在许多复杂结构的形成中都扮演了重要角色。选择性修剪是发育和进化的核心特征，但并不是说身体在生长过程中进化出最终的形态。进化的思想不仅仅是适者生存。复杂引擎的核心是有细微变化的复制，然后选择最符合某种标准的复制体。发育过程中的选择通常不是因为基因变异，而是由于细胞特性与环境信号不匹配。细胞在不同的微环境中可能被激励或压制，取决于它们所处的环境。在这里细胞中的基因变异不起作用，被选择的结构也不会复制。

为了说明（无复制的）构造引擎与（复制的）复杂引擎的相似和不同之处，这一章将对大脑的形成与免疫反应进行比较。这两个系统都过于复杂，无法用细胞中的DNA直接描述；并且对于这两个系统，每个人都表现出独特的细胞结构、特征和分子层面的作用。两个系统的独特性都源自累积性的随机选择，但两个系统中随机性的产生和使用方式不同，只有免疫系统才具有对多重复制的选择。

大脑是如何连线的？

人类大脑有1000多亿（10^{11}）个神经元，与手的细胞数量差不多；但每个脑细胞平均有10000（10^4）条细胞间的连接，远多于手细胞。这意味着大脑有大约1000万亿（10^{15}）条细胞间的连接。这些连接的模式决定了我们是谁，怎么做，怎么想，以及记得什么。人类DNA有大约30亿（3×10^9）个核苷酸。根据DNA测序的结果，目前估计的基因数量介于20000（2×10^4）到30000（3×10^4）之间。如果所有人类基因都参与脑结构的形成，每个基因就要负责说明1000亿（10^{11}）条细胞间的连接，这当然是不可能的。如果DNA中所有核苷酸都用于说明神经连接，

则每个核苷酸要说明300000条细胞间连接。这个结论显然很荒谬。因此可以肯定神经元之间的连接不是由基因直接说明的。然而，我们都拥有人类大脑而不是狗的大脑或昆虫的脑，是因为我们有人类的DNA。这导致了概念上的两难境地。

要理解如此复杂的功能如何在没有明确指令的情况下产生出来，要考虑大脑的几个特点。首先，虽然所有正常的人脑都大致类似，都有相同的解剖特征，都有皮质、下丘脑等，连接各部分的神经传导束也非常相似，在显微解剖细节上每个人的脑却大不一样。神经元之间的连接细节因人而异，就连同卵双胞胎的也不一样。似乎DNA的信息只说明了大脑发育的概略或总体规则，而没有具体的细节。现代神经科学的一个主要挑战是，理解大脑的设计——在没有详细的连接说明的情况下——如何实现可靠、复杂而且具有目的性的功能。这种设计策略完全不同于制造计算机时采用的策略，制造计算机时同一个型号的每台产品在微观细节上完全一样。如果没有这种明确的设计，电子计算机就无法工作。只有对大脑连接的粗略说明又是如何创造出像人脑这样复杂的信息处理结构的呢？几乎可以肯定答案就是选择。

杰拉尔德·埃德尔曼提出，一种他称之为“神经元群选择”的进化范式对于理解大脑很关键[2]。他认为，在初级大脑的形成过程中利用了进化原理，而且进化原理在学习中也扮演了重要角色。在第10章我们将看到，在从信息的角度分析学习行为时，产生了另一个哲学困境，而复杂引擎则提供了可能的解答。

这一节关注的不是学习，而是作为学习基础的神经元之间的初级连接的结构形成——也就是大脑如何连线。我们将看到，修剪掉不想要或不必要的连接扮演了关键角色。大尺度脑特征的形成面临着与肝脏或手等其他身体结构的形成同样的难题，但神经元连接的具体细节带来了另一个复杂性层面。毕竟，肝脏细胞与其他哪些肝脏细胞相连无关紧要，而脑细胞与哪些脑细胞相连则至关重要。

一旦细胞根据设定的程序迁移到大致合适的环境，并消除掉错位的细胞，胚胎神经元就会停止迁移，并受长程信号吸引，开始伸展出又长又细的突出部分。在细胞伸展的过程中，它们会探索经过之处的局部环境。一旦到达合适的区域，就会在细胞之间形成称为突触的连接。如果遇到了不合适的细胞，探索过程就会回退并重新定向到其他地方。不是

所有突触都有作用。如果形成的突触有作用，伸展的突出部分就会被加强；如果形成的突触不起作用，就会消退。如果神经元没有形成任何有用的突触，就会死亡。长程吸引和抑制信号引导神经元的生长，将探索聚焦于特定的远距离细胞群。通过这种方式，神经（细胞扩展束）投射到大脑和身体的不同部位，并与远距离的神经节和肌肉等目标形成功能连接。这种自我连线活动虽然由分子信号引导，却有可能受噪声影响，导致无法连接和出现不正确的连接。这个过程的一个关键是加强正确的连接和消除不正确的连接。连接的选择基于简单的规则，规则表现为每个细胞表达的特定蛋白质的结构，当然，这些蛋白质结构是由DNA的语言写成的复杂规则决定的。

神经元伸展出来与其他细胞连接的部分被称为树突和轴突，树突从其他细胞接收信号，轴突则向其他细胞发送信号。图7.2展示了几个例子。树突的形状很像树，也因此而得名。图中没有显示与其他细胞连接的精细分支。突触位于轴突的顶端，并沿树突的端沿分布。这些结构形成的时间和地点受约束，但并没有预先确定细节。大脑特定区域特定类型的神经元一般外形相似，但每个细胞具体的细节都不一样。相似的地方包括突触数量大致相同，树突的形状大致一样，连接的细胞类型也一样。连接到哪些细胞的哪个部分则似乎有很大的自由度。伸展的部位一旦到达大致的目的地，就会不断分叉探索局部区域；一些分叉成功连接，没有成功的则缩回。最终的树突形状由基因、随机的局部探索以及环境细节共同决定。由于细胞层面的细节各不相同，以及探索的随机性，最终的结果也就因人而异。

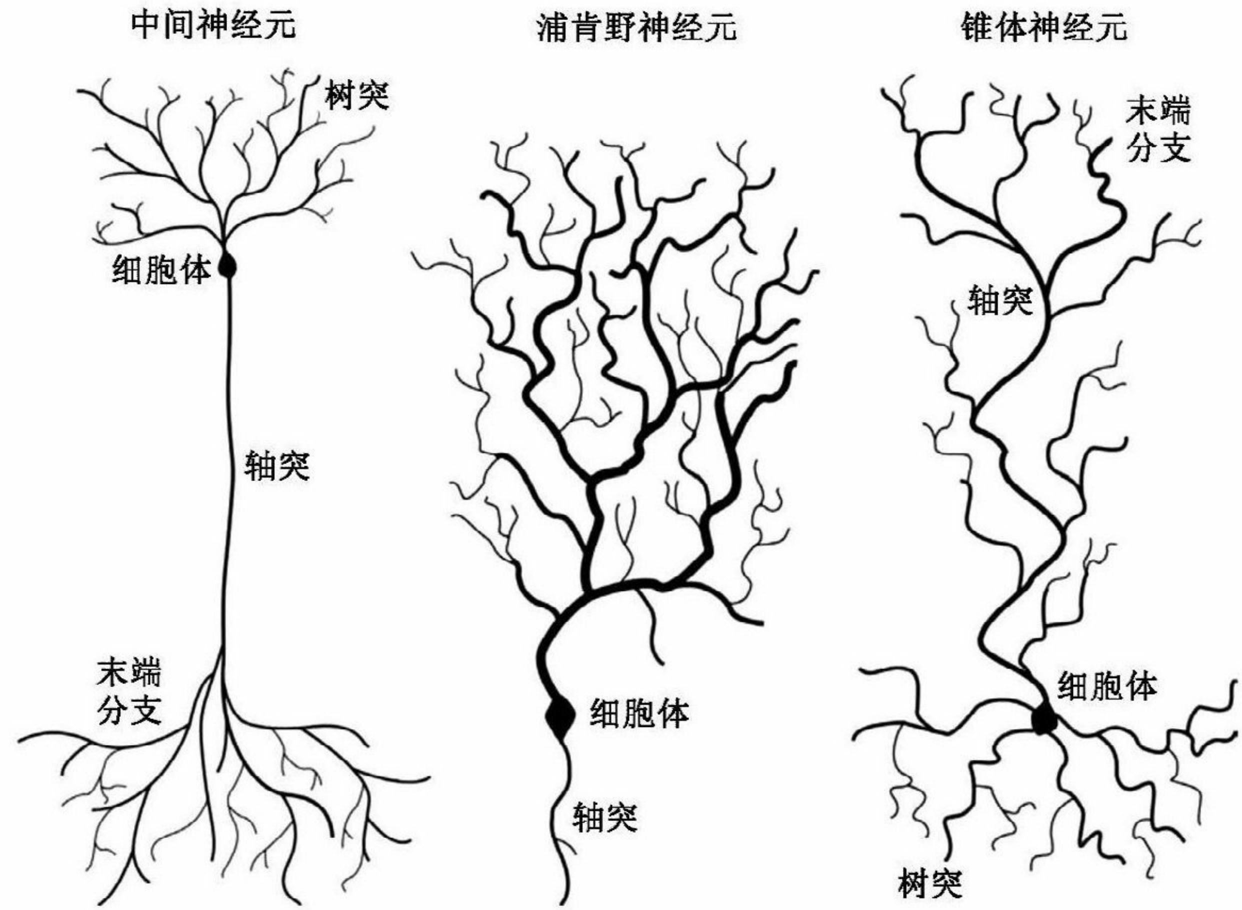


图7.2　由细胞体、轴突和树突构成的神经元。纤细的轴突和树突都是细胞的分支扩展，与其他细胞形成连接（突触）。树突从其他细胞收集信息，轴突则向其他细胞发送信息。每个神经元的具体分支模式都是独一无二的，只是特定类型的神经元相互类似

在神经元连线的发育阶段，历史偶然性起了重要作用。当分支形成，如果受到刺激就会增强，如果没有刺激就会消失。显然一旦分支消失，就无法作为进一步分支的基础。通过这种方式，树突的形状以3维形式编码了过去分支决定的记忆。每个神经元都通过随机的分支发动、半随机的环境探索以及对产生的分支的选择，形成了独一无二的结构。

树木的生长也采用了非常类似的原理。遗传上一样的榆树具有相似的总体形状，但具体的分支模式每棵树都不一样。分子和细胞层面的随机事件通过与环境的互动被选择，最终的结果是独一无二的3维结构。脑神经元的分支不是伸展到空气中，而是其他细胞组成的3维环境中。连接细节由分支与分子信号的互动以及是否建立起突触决定。分支的生长受内部和外部信号刺激，但分支过程显然是随机的，创造非随机的连接模式则由随后的选择负责。

总结一下这一节，脑的发育受微观上随机的探索以及对不正确的连接的修剪引导。神经元的分支很相似，因此可能会被认为是“复制体”，但神经元分支的具体细节并不能在其他地方复制。结构必须重新构建，不受过去的经验影响。因此，脑的发育并没有利用复杂引擎的核心特征，虽然脑的发育很惊人，但采用的并不是反复选择的策略。身体中还有一个系统的确结合了选择与不完美的复制，就是免疫系统。

为什么身体没有被微生物摧毁？

病毒和细菌会让我们的身体生病，这些微生物只需几个小时就能在体内大量繁殖。进化是以代为时间单位，因此微生物的进化速度比我们快上万倍。我们花上百万年才能进化出的分子防御策略可能几年甚至几个月就会被对手进化出的策略破解。那为什么我们没有经常因为微生物感染而生病呢？简单说是因为对于大部分微生物，我们体内的免疫系统能针对每次入侵迅速进化出分子响应策略[3]。

要防御这些入侵者就需要能识别一些体内通常没有的分子（蛋白质、多糖、脂类）。防御系统通常分为两部分，先天免疫和获得性免疫。先天免疫包含多种机制，利用入侵者的普遍特征识别和消灭它们。例如巨噬细胞会在体内巡逻，寻找和吞噬不应该存在的东西。但是一些可能的入侵者进化出了能骗过先天免疫系统的办法。这些更具挑战性的入侵者由获得性免疫系统处理。

获得性免疫系统依赖于特异性识别分子的识别；所有生物都有获得性免疫。这种系统有两种机制：一种称为体液免疫，由B细胞实现；另一种称为细胞免疫，由T细胞实现。两类细胞都产生特异性蛋白质，能专门与某种分子结合。B细胞能产生名为抗体的蛋白质，而T细胞则能产生一类名为T细胞受体（TCR）的蛋白质。抗体通过与体液中的外来大分子（蛋白质或多糖）结合来对其进行标记，而T细胞则通过细胞体上的TCR与外来蛋白质的片段结合。两者都能“记住”之前识别的外来分子，并对以后同样分子特征的入侵快速响应，这个记忆就是我们所说的免疫。

B细胞和T细胞都是对被称为抗原的蛋白质或多糖的特异性分子尺度特征进行响应。由于你自身也产生蛋白质和多糖，因此这个系统要能

够识别外来抗原，而不是自身抗原。在分子尺度上，“识别”意味着结合。当抗体识别出抗原，就会与其结合，从而阻止其行动，并引起免疫系统的注意。

蛋白质和多糖是复杂的大分子，在自然界中种类很多。这意味着适应性免疫系统可能遇到的抗原种类极多，远远多于大脑中的连接数量。因此免疫学有一个核心问题是，外来蛋白质和多糖的结构类型多得不计其数，免疫系统是如何产生出能对任何可能的外来抗原进行特异性识别的抗体（和T细胞受体）的？

过去30年的研究揭示了部分答案，就是你的身体能产生大量不同的抗体。在任何时候，你的血液中都包含数以百万的抗体，各有不同的结合特异性。而你的DNA并没有编码这数百万种不同抗体的基因，要理解你的身体是如何产生出能与从未遇到过的抗原进行特异性结合的抗体，我们需要了解抗体的结构和产生抗体的细胞的性质。

免疫系统是由散布在整个身体中的白细胞组成的分散器官。B细胞、T细胞和巨噬细胞（变形虫状细胞）都是白细胞，还包括其他一些细胞类型。每种细胞类型都在身体防护中扮演各自的角色。在尝试理解免疫的分子基础时，要记住一个重要的原则，就是每个B细胞都只能产生一种抗体的许多拷贝（即相同蛋白质设计的许多副本）。每个T细胞也只能产生一种T细胞受体。这意味着有多少不同的B细胞和T细胞就有多少不同的抗体和TCR。我们重点关注B细胞。

所有抗体的结构都包含一个分子面，有可能与某种特异性的抗原结合，抗体和抗原的表面精确匹配，就好像一把钥匙开一把锁。抗体的抗原结合面各不相同，决定了抗体的特异性。系统作为整体要能识别以前从未遇到过的外来抗原，就必须产生出大量不同的抗体，各有不同的抗原结合位。由于每个B细胞只能产生一种抗体设计，因此必须产生大量不同的B细胞，各对应不同的抗体结构。主干策略很简单。如果产生大量不同的B细胞（以及相应的抗体），那么出于随机其中一些将有可能识别出外来入侵者，与其表面的某种抗原结合。由于外来入侵者是蛋白质等生物分子组成的生命物质，它们必须与外界互动，也就不可避免地具有某种独特的表面分子特征，因此就有可能被某种抗体识别。

由于抗体的结构是由氨基酸序列决定，而氨基酸序列又是由DNA的

核苷酸序列决定（回想一下第4章讨论的蛋白质的构成），因此不同B细胞的DNA（抗体基因）有不同的核苷酸序列。但这怎么可能呢？每个生物系的新生都会被教导说，我们所有细胞中的DNA都是一样的。答案是这个简单的信条对产生抗体的细胞不成立。皮肤或肝脏细胞（任何非免疫系统细胞）没有抗体基因，只有基因片段。只有B细胞拥有完整的抗体基因。在B细胞成熟过程中，基因片段（随机）组合成新的抗体基因。这是为什么在同一个身体中能产生如此多不同的抗体蛋白质序列的第一层解释。每个B细胞都有独一无二的抗体基因。

抗体由4个蛋白质亚基（4条氨基酸链）组成，两个大亚基称为重链，两个小亚基称为轻链。两条重链和两条轻链分别是一样的。因此，所有抗体都是由两个基因决定，一个负责重链，一个负责轻链。所以每个B细胞都有两个独特的基因负责产生抗体。图7.3展示了抗体的分子结构。一个B细胞可以合成许多相同的重链和轻链，这两种链又会组合成许多相同的抗体。

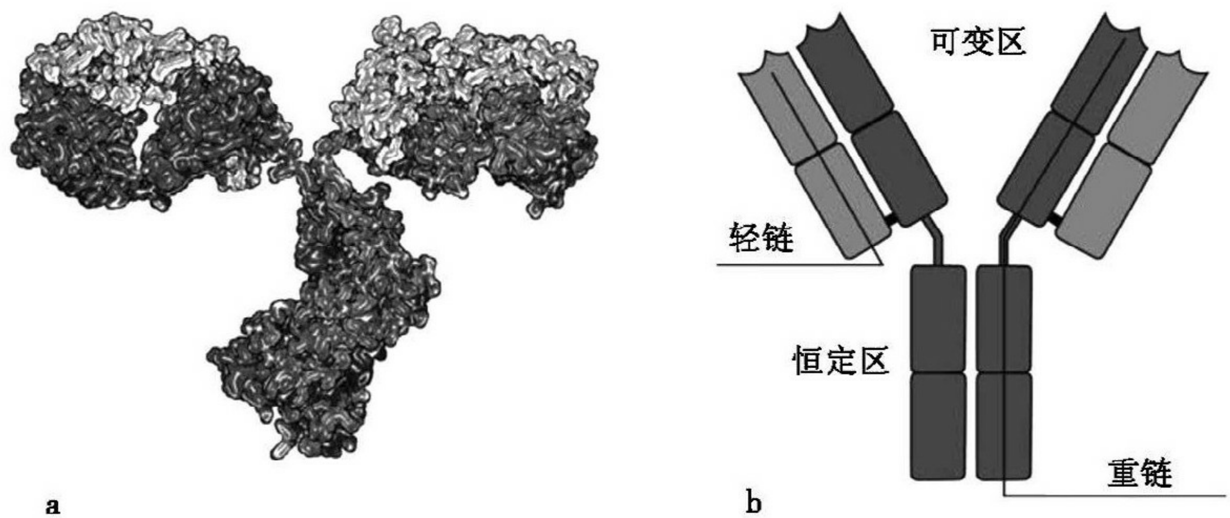


图7.3　a：抗体原子级的3维空间填充模型（RSCB蛋白质数据银行版权所有）。b：抗体结构原理框图。抗原在顶端的可变区结合

对许多抗体的物理和化学分析表明它们都有恒定区和可变区。一类抗体的恒定区都是一样的，而可变区则各不相同。与抗原结合的是可变区。可变区既有重链也有轻链。由于特定的B细胞产生的重链和轻链都是一样的，因此一个细胞产生的所有抗体都能与相同的抗原结合。

B细胞在骨髓中成熟的过程中，每个新生B细胞的DNA都会重排，

从基因片段合成一个重链基因和一个轻链基因。轻链可变区片段有2种类型，V和J；而重链可变区片段则有3种类型（V、J和D）。人类非B细胞的DNA包含70个功能轻链V片段和9个J片段。哪个V段与哪个J段结合是随机的，只有重链和轻链的DNA功能性重排都完成后，B细胞才能产生抗体。这些DNA重排能产生320种不同的轻链和6000种不同的重链。轻链（320种可能）和重链（6000种可能）的随机匹配，能产生190万种不同的重链和轻链组合。这意味着DNA重排就能直接产生大约200万种不同的抗体基因。这个数量很惊人，但与实际观察到的抗体多样性比起来还只占一小部分。

抗体多样性的部分原因在于B细胞的DNA重排不是很精确。甚至可以说是相当马虎。在V区和J区结合之前，特定的酶会在片段端部随机添加和移除核苷酸。由于DNA核苷酸的序列决定了组成抗体的氨基酸序列，因此V和J结合部的核苷酸序列的随机修改使得产生的抗体种类比DNA精确结合能产生的种类多得多。这种有意的马虎机制付出的代价是产生的抗体基因中许多（超过2/3）都不具有功能。编码了非功能抗体的B细胞有两种命运：要么再进行第二次DNA重排，要么死去。

如果一个B细胞碰巧能识别某个外来抗原，则会带来更多变化。骨髓中的干细胞每天会产生大量不成熟的B细胞。这些不成熟的B细胞会通过DNA重排形成抗体基因。一旦重排完成，新合成的抗体就会出现在细胞表面。这些细胞的命运由抗体结合的东西决定。如果与骨髓中存在的抗原结合（自身抗原），细胞就会死亡或变得无反应。这些细胞被从B细胞群体中移除。这称为克隆删除，因为它们没有后代。没有与自身抗原结合的B细胞产生的抗体则被释放到血液中。

B细胞的生命期很短，只有几天，除非其表面的抗体与外来抗原结合。如果是这样，外来因子的存在就会被同时的T细胞活动确认（后面讨论），然后B细胞被激活。激活导致细胞的一系列变化。它开始分裂，超突变被激活，之前附在细胞表面的抗体被释放到血液和淋巴中。最终产生两种后裔：分泌大量自由抗体的原始细胞和不断在表面表达抗体的长寿命记忆细胞。记忆B细胞为长期持续的体液免疫提供了基础。在被之前遇到过的疾病再次感染时，这些细胞会被迅速激活。

体液免疫的一个困扰多年的问题是感染后产生的抗体会比感染初期产生的抗体更具特异性，与外来抗原的结合更为紧密。现在知道了这是

因为超突变和选择。一旦B细胞被激活，就会复制其DNA准备细胞分裂。产生特定的酶让可变区抗体基因DNA的复制变得容易出错。其他DNA则不会受这些酶影响。重链和轻链基因可变区核苷酸序列的错误率估计约为每一千个核苷酸错一个（10^{-3}），而常规DNA复制的错误率约为每1亿个错一个（10^{-8}）。综合重链和轻链可变区的大小，大约每两次细胞分裂，重链和轻链DNA的复制就会发生一次错误。每次错误都会导致抗体的变化。由于变化发生在可变区，抗体也就具有了不同的结合特异性。大部分结合都不怎么样，但一些会与目标抗原结合得更好。被激活的B细胞的命运由其抗体结合的紧密程度决定。

抗体的结合并不是要么有要么无的关系。一些抗体不止与一种抗原结合，任何一种抗原都可以被多种不同抗体识别。一些抗体抗原反应较强，一些则较弱。B细胞只有在抗体与一种外来抗原有至少中等强度的结合时才会被激活。在B细胞克隆扩增时（快速细胞分裂时），B细胞合成抗体变体的超突变类似又不同于能识别外来抗原的原抗体。大部分变体的结合都不如原抗体，这样的B细胞不会再分裂。但一些变体会比原抗体结合得更好，它们也会因此分裂得更快。很快结合得更好的B细胞就会超过结合得不好的细胞。超突变贯穿B细胞的整个扩增过程。因此，经过几天或几周，随机突变就会发现越来越好的抗体，与之相应的B细胞也会比抗体结合较弱的细胞更容易分裂。产生的抗体与外来抗原结合最紧密的B细胞分裂也最快。这些B细胞在血液中会占主导，它们产生的抗体成为血液中最常见的抗体。这个受限的随机修改然后选择的系统会产生并不断改进抗体。

超突变和克隆选择组成了身体里的进化。最适应的B细胞生存，适应性则是由抗体与存在的外来抗原的结合紧密程度决定。利用这种简单的方法，对于几乎任何分子结构，身体都能产生出能紧密结合的抗体，因为对每种新的抗原都能进化出适用的抗体。虽然感染会造成一时的问题，但每种新的感染都会导致进化出新的抗体来识别和摧毁病原体。

第5章列出的进化过程的所有要素（参见“指令从何而来”一节），以及图5.2中展示的复杂引擎的所有特征在免疫系统中都有呈现。可遗传性是基于DNA，细胞群体进化，遗传信息产生随机变异，特定序列的细胞被选择生存并繁殖，其他细胞则消亡。免疫反应的确体现了身体中的进化。对入侵微生物具有高度特异性的新细胞一两周就能进化出来。

B细胞和它们产生的抗体只是免疫系统的一个方面。抗体包裹外来抗原的表面，干扰其功能，并对其进行标记，让在体内巡逻的巨噬细胞将其摧毁。由于抗体结合的效果非常强，如果产生的抗体与身体自身的细胞和蛋白质结合会产生很大损害。作为防范，B细胞的激活（通常）需要T细胞的协同刺激。刺激B细胞所需的双重信号（抗体结合与T细胞刺激）确保不会产生大量自身免疫反应的抗体。

T细胞有多种亚型，包括T辅助细胞和T杀伤细胞；类似于B细胞，T细胞在表面会产生名为T细胞受体（TCR）的特异性蛋白质。TCR也能识别外来抗原。类似于抗体，T细胞受体也具有可变区和恒定区，可变区可以与抗原结合。同样，类似于抗体，非T细胞的DNA没有完整的TCR基因，TCR基因也是由片段组装，方式类似于抗体基因。

T细胞也是通过造血干细胞分裂从骨髓中产生。不成熟的T细胞转移到胸腺进行DNA重排，用V、J和C基因片段产生功能性TCR基因。类似于抗体，在结合过程中，结合区的核苷酸会随机增加和删除，从而使得TCR具有极大的多样性，各具不同的特异性，同样类似于B细胞，成熟T细胞是从不成熟T细胞中选择出来的。

TCR识别的抗原与抗体识别的不同。在产生的数百万抗体中，一些能与适当大小的任何分子结合，而TCR则只与同时与细胞表面的MHC蛋白质结合的多肽（氨基酸短链）结合[4]。体内几乎所有细胞都会分解一小部分其生成的蛋白质，将产生的片段组合成MHC蛋白质。肽MHC复合物被输送到细胞表面。因此体内大部分细胞都会被其产生的蛋白质片段修饰。

T细胞不断检查这种缩氨酸。一旦发现是外来抗原，下一步发生的事情就取决于目标细胞的性质。对于大多数细胞，外来缩氨酸抗原的出现给出了一个信号，表明这个细胞被病毒或细菌感染了。通常，这种细胞会被杀死，摧毁入侵者的藏身之所。B细胞的行为不同。不成熟的B细胞的表面有抗体。当这种表面抗体与外来抗原结合，抗原会被包围分解，产生的片段会出现在细胞表面。如果“呈现的”缩氨酸片段被T细胞识别为外来物，B细胞就会收到第二个信号（与抗原结合是第一个信号），确认其确实遇到了外来物质。通常需要这种双重刺激，B细胞才能被激活进行快速细胞分裂，并成熟为分泌大量抗体的原始细胞。

由于T细胞调控免疫反应，因此T细胞不识别自身抗原非常重要。发育的T细胞在成熟过程中也会经历激烈的选择清除阶段。被选择存活的细胞从胸腺中释放出来，迁移到胰腺和淋巴结中。TCR基因的形成没有超突变阶段，但在结合过程中有广泛的随机变化。总体效果类似于B细胞的克隆选择。

同B细胞一样，对外来抗原具有高度特异性的T细胞的产物会影响大量具有不同TCR的细胞的繁殖，通过强烈的反向选择，结合特征不理想的细胞被清除，能识别不应存在的蛋白质的少量细胞则被刺激快速增殖。B细胞的成熟过程则更进一步，利用了反复的突变和选择过程，针对每种入侵的微生物进化出不断改进的特异性反应，速度比大部分微生物进化出绕过防御的机制的速度更快。B细胞的反应充分利用了复杂引擎，这解释了为什么它能如此精妙地应对各种新的疾病。

复制对选择有何助益？

在从单个细胞发育成人体的过程中，充分利用了反复的随机变异与非随机的选择。从胚胎到成人的身体发育过程的所有阶段都表现出了细胞层面的随机活动，并且身体的形成完全依赖于这些活动带来的新的可能性以及对其中满足预设标准的结果的选择。新的结构总是基于已有的结构，并且在所有阶段选择都亲睐其中某些细胞，排斥其他的。留存下来的细胞成为新细胞的父代，新细胞又接受新一轮的选择。除了免疫系统，细胞的发育和细胞的扩增都不是基于遗传选择，而是基于位置的选择。位于正确位置的有正确连接的细胞被激励，否则被清除。这个过程很像雕塑。这个策略的一个自然结果是个体之间在细胞层面和分子层面上的多变性，虽然在大的尺度上很相似。

新的B细胞的产生与这个策略有一点很不一样。在大脑发育过程中，编码产生神经元突起的蛋白质的DNA序列并不变化。当轴突和树突生长探索特定的邻居时，基因表达的模式改变，但背后的指令并不变。DNA引导有利过程的生长，同时引导不利过程的消亡。而在免疫反应中，特定细胞的DNA血统却会改变。对于B细胞，编码抗体可变区的DNA序列不断变化，每个细胞的血统都是独特的。通过允许指令变化，系统具有了适应任何可能入侵者的最大潜力。大脑的发育也很惊人，但表现出的创造性要有限得多。如果大脑发育过程中对神经元也有遗传性

选择，则人类大脑将不仅在微观层面上不一样，而是在所有方面都会不一样。如果是这样，大脑将不是依物种而异！这样大范围的变化在物种进化过程中（经过数百万年）的确会发生，但在个体发育过程中不会发生。

免疫系统应对外来抗原发生的变化类似于生物进化，但是在体内很短的几周时间就会发生。身体不可能预先“知道”下次会遇到哪种细菌或病毒袭击。

对随机的微小变化（突变）进行复制，然后选择表现好的变体，这为解决面临的问题（发现新的蛋白质结构）提供了无尽的可能。快速进化的微生物有各种编程的可能性。超突变结合克隆选择确保了一旦弱结合的B细胞被激发，很快就能产生出具有高度特异性的抗体。

基于线性编码信息的计算的用途如此广泛而强大的原因之一是很容易复制。复制1维和2维对象很容易，复制3维对象则很难。原因在于几何。这也是为什么你能买到复印机却买不到《星际迷航》中的那种复制机。如果某种东西有用，则往往是多多益善。并且稍加修改，经常还能变得更有用。物理世界的大部分事物都是3维的，无法（直接）复制，但1维指令可以用来产生相同3维设计的多个对象，指令为绕开复制3维事物的难题提供了途径。制造也利用了相同的原理。不那么明显的一面则是对指令的细微修改往往能得到新的有用的事物。这一点从根本上使得产品开发、动植物进化和适应性免疫成为可能。

通过大规模并行利用抗体结构的随机而细微的变化，身体得以产生合适的抗体对抗蛋白质组成的任何入侵者。由于地球上所有生命都是基于蛋白质，这是一个很可靠的策略。在大脑发育过程中，随机连接的形成为探索复杂的局部环境提供了高效的途径。但与生成抗体的情况不同，产生轴突和树突所需的DNA信息在人的一生中并不改变。通过对不符合预定标准的神经元连接进行反向选择，就可以从初始的混沌状态中构建出极为复杂而又多少可以复制的连线模式，并且无需预先存储大量信息。这种策略的一个结果是，大脑不像买来的计算机那样具有“现成的”功能，要经过学习才能发挥作用。在第10章我们将对此进一步探讨。

第8章　控制循环

复杂引擎能接受任意形式的信息吗？

人类身体和大脑的发育过程很神奇。它证明了物理系统的创造性潜力，产生变体，进行选择，并基于选择创建新的结构。但这种策略也有局限。在物理世界有许多事物的实现并不是这样简单。这就是为什么指令在我们的世界中如此重要。从本质上来说，只要有合适的指令，物理世界的一切都是可实现的。复杂引擎在指令层面运作。将选择变体的基本原理应用于指令，就有可能产生出新的指令，从而开启无限可能。前面我们探讨了3个例子：生命，针对疾病的免疫，以及进化出难题答案的计算机程序。在后面的章节中我们还会看到技术和人类思维也是同样的计算策略的产物。但在这一章我们将探讨的问题是，在多大程度上复杂引擎可以被人类工程师利用、控制和操纵，从而为人类造福。

所有复杂引擎计算都有3个基本要素：①以可复制形式编码的信息体；②引入有限变化（经常就是复制时的错误）的复制机制；③决定哪些复制体被再次复制的选择方法（图5.2）。我们要问的是有没有什么信息结构不适合用复杂引擎进行修改，其实也就是问有没有什么结构无法被复制。

所有生物都是用细胞内部的信息来编码说明其结构的许多方面。这种信息表示为组成DNA分子的核苷酸序列。当细胞繁殖时，DNA及其编码的信息就会被复制。这个系统能编码很多信息，因为DNA分子（作为分子来说）很大，可能的核苷酸序列极多。同样的编码系统也为我们的适应性免疫系统提供了基础，只是采用了不同的进化方式。计算机算法采用各种顺序结构编码信息，但最终都可以归结为0/1序列。人类思想和文化也是基于编码信息，这里信息被编码为大脑中的神经元连接模式。

前面我们看到，只要系统（有可能）有多种形态，就有信息与各种形态相关联。信息在生物界和人类文化中之所以如此重要是因为它能表示某种东西。这带来了意义。通过给特定的形态赋予特定的意义，就有

可能为构造某种东西编码指令。观察到的信息形态的多样性带来了一个问题，就是有没有哪种信息不适用于基于复杂引擎的变化。复制环节似乎的确施加了限制。几何很重要。一般来说，1维和2维结构更容易复制，3维结构则有问题。

至少有3种方式可以实现复制。可以直接复制，例如T复制为T，A复制为A，可以利用中间媒介，也可以利用指令。直接复制在自然界和技术上都不常见。印刷机利用模板；印出来的每一页都是基于一个特殊的滚筒生成的镜像产生。后面我们会看到，DNA复制时，每一条链都是新链的模板。旧链上的A核苷酸引导在新链上组合一个T核苷酸，同样，T引导A，G引导C，C引导G。用这种方法，新链的序列总是由模板链的序列决定。工业生产也经常用模具生成想要的形状。模具可以复制简单的3维形状，例如塑料瓶或青铜雕像。模具不适用于复杂的3维形状。复杂对象的复制不能直接进行，用模板也不行。要复制这样的对象，首先必须得到3维对象的1维或2维表示，然后用对象的表示来引导生成3维对象副本的过程。通过对表示进行复制可以扩展过程。指令是可以用于生成3维对象副本的1维表示方法；蓝图则是2维表示。

前面曾讨论过蛋白质的结构。蛋白质在细胞和身体的功能中扮演了关键角色。没有蛋白质就不会有生命。蛋白质的功能由形状决定。在复制蛋白质时，细胞以信使RNA的形式生成1维表示，然后将这种表示转录成氨基酸链，氨基酸链自发折叠成所期望形式的新的拷贝。信使RNA序列总是生成同样形状的蛋白质。当DNA被复制，信使RNA合成，蛋白质的表示——即基因序列——就被复制了，但折叠的蛋白质不会直接以3维形式复制。

产品的生产也很类似。汽车不是直接从另一辆汽车复制；而是基于同样的设计生产同样的汽车。只要设计一样，生产出来的汽车就一样。如果要改进，就要修改设计，然后制造新的模具。这就是指令的魔力；它们使得不适合直接或通过模板复制的复杂（3维）结构的复制成为可能。将结构表示为1维的指令形式使得任何能建造出来的对象都能被复制。

由于3维形状总是可以表示为1维编码的信息，原则上1维表示（指令）总是可以被复制和修改，然后变换成新的3维对象。这个变换、修改、再变换的循环是生命和技术的基础。因此这一节的问题的答案是

否，不是所有信息形式都能直接或通过模板复制，不能这样复制的也就不能作为复杂引擎的输入。但是，似乎所有结构都能表示成1维编码，可以被复制、操作，然后用来生产3维结构的新的“复制体”。因此，人体不能被复制，但人类DNA可以，而人类DNA一旦被放置到适当的环境（人类受精卵）中，就能（通过发育过程）产生新的人体。通过复制DNA并在合适的细胞环境中激活正确的基因顺序表示，原则上就有可能创造出同一个人的多个复制体。好莱坞称之为克隆人。用人类DNA进行这样的实验目前在全世界大部分国家都是违法的，但其他生物DNA的实验很常见。

变化可控吗？

复杂引擎是一种信息操作策略，创造变体，然后非随机地挑选。输出结果的可变性为从中选择提供了可能。在复制过程中引入变化的最简单方式是允许发生错误。在分子层面上，错误通常是随机的。因为变化是分子热运动造成的，因此不太可能是为了系统未来的某种需求。但这并不意味着随机变化对引擎是必需的。如果复制过程中引入的变化不是随机的，复杂引擎仍然能够运作得很好。任何来源的变化都能为选择提供材料。然而让人吃惊的是，随机变化能够提供最大的潜力。如果变化不是随机的，系统也能进化，但未来的潜力会受到限制。如果对未来应当是怎样有预先的构想，则所有没有想到的可能性都被排除。只有变化包括随机因素，才为系统真正打开了机遇之门。随机变化将未来的潜力最大化。

抽象地思考信息的变化很难，我们可以看看非常重要的进化系统——生命——引入变化的各种方法的一些细节。生物新的可遗传变化的最终来源是错误和灾难，但生命利用了一些窍门来充分开发这种资源。复制DNA的分子机制很精巧，但并不能排除错误。DNA分子很长而且易断，许多化学物质和大部分高能辐射都能导致DNA受损。因此你的DNA中编码的信息不断受到攻击。细胞可以产生酶来探测和修复损伤，但修复机制不是完美的。这实际上有好处；错误对进化的不断进行很重要。如果某个物种发展出了能完美复制和修复DNA的机制，未来就无法再进化。最终，环境会改变，其他进化的生物会改变以适应新的环境，并在对食物等资源的竞争中胜出。之前完美的生物就会越来越处于劣势并走向灭绝。没有错误，就不会有变化。在生物界，不变的完美是灭绝

的保证。

错误的关键作用导致了一个惊人的结果，就是经过数百万代后，生物已经进化出了它们的突变率。复杂引擎只在很窄的一个突变率范围内才能最有效地运作，变化率要足够大，但又不能过度，以至于丢失之前积累的有用信息。变化太多同太少一样有害。图6.4说明了这一点，生成100位1字符串的最多1算法的最优突变率是每一代每一位2%。在自然界观察到的突变率也说明了这一点。基因组包含10000（10^4 ）个核苷酸的病毒的突变率约为每一代每个核苷酸10^{-4} 次突变。而有400万（4×10^6 ）个核苷酸的细菌的突变率接近10^{-6} ，DNA中有30亿（3×10^9 ）个核苷酸的哺乳动物的突变率约为10^{-8} 。

DNA还有一种对生命进化很重要的变化机制是保留副本。当一个或一段基因在同一个DNA分子内部有副本时，就不容易损失信息。

当一份拷贝在未来变化时，副本会保留原来的功能。对动植物基因组的分析发现进化变化产生的副本使得细胞蛋白质的扩充和多样化成为可能。对序列的删除则是另一种更具破坏性的变化来源。

还有两种相当不同的机制确保最初由突变、副本和删除引起的变化能与其他基因活动“适配”并促进有用的新序列的推广应用。这两种机制是减数分裂过程中染色体的随机配对和DNA的重组，这个过程产生精子和卵子。这些机制很重要，因为不仅仅是外部环境决定DNA的变体是传播还是消失，生物体内的基因（以及基因变体）的互动也起作用。在有性繁殖的背景下，这些机制确保每个后代都能获得其父母完整而又独特的基因变体的混合。重组使得信息可以混合，随机配对则使得每个后代都能拥有4个祖父母的基因变体[1] 。

染色体的随机配对会混合群体中已经出现的序列变体。所有多细胞生物的细胞中都有多个DNA分子；DNA分子被蛋白质和RNA包裹，形成的微观结构就是染色体。DNA分子短，染色体也短；分子长，染色体就长。由于细胞中DNA分子长短不一，因此实验人员可以根据染色体的长短来对其进行区分。另外，大部分多细胞生物的每个细胞中都有每个DNA分子的两个不同备份。这意味着每个细胞中每种染色体都有两条。每一对有一条来自父亲，另一条来自母亲。人类细胞有46条染色体，两条一对，共23对。豚鼠细胞有64条染色体，果蝇有8条，玉蜀黍有20

条。每条染色体中的DNA编码一组特定的基因。在有性繁殖过程中，配偶子含有每对染色体中的一条，但具体哪一条是随机的。因此，来自每个祖父母的染色体平均各占1/4。但只是平均；具体的分布是随机的。

图8.1展现了重组。最常见的形式是同源重组，两个包含相似但不完全一样的序列的DNA分子（图中的1和2）对齐，使得相似的序列并列；然后两个分子都在随机选择的同样位置截断并交错重连，这样截断点两边各来自父母的一方。

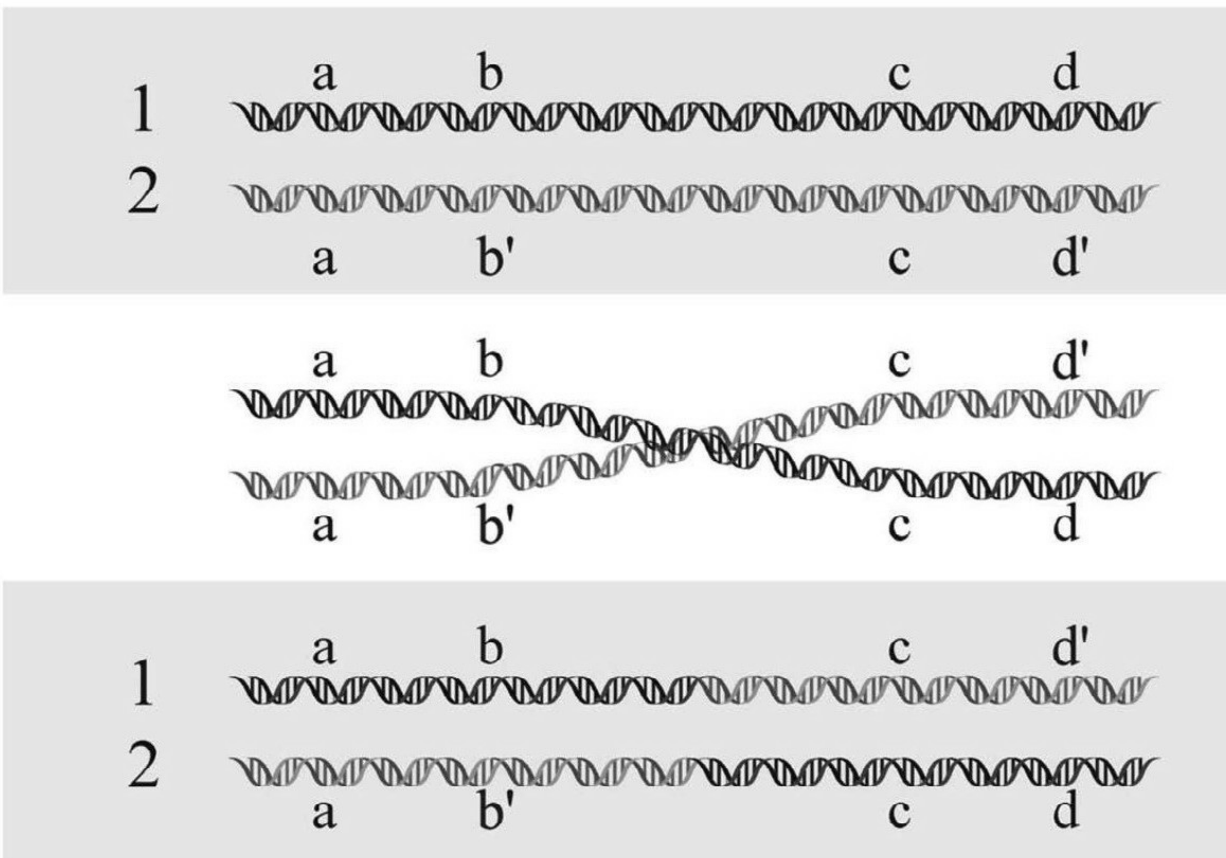


图8.1　每一条横线代表一个DNA分子（来自父母中的一方），线上的字母代表核苷酸序列编码的一个基因。由于大多数生物都有两个父母，因此每个基因都有两份拷贝，这两份拷贝稍有区别，用（′）标注。在重组时，携带同样基因的两条DNA链排到一起，在相同的位置断开，然后交错重连

随机配对和重组的组合保证了卵子和精子结合产生的后代都有独特的基因变体组合，而且每种基因都有两个。这反过来又保证了每个基因变体都有与其他基因变体组合的机会。在许多遗传背景下具有优势的DNA序列将有可能在群体中广泛传播；没有优势的则会消失。复制和修

复，以及更为罕见的副本和删除导致的错误使得新的变体不断产生。如果新变体的引入太过迅速，系统就会被错误占据，并因为过于缺乏（有用的）信息而崩溃。如果引入的速度太低，变化就会非常缓慢。

表8.1列出了生物系统变化的主要来源。分子生物学的最新进展带来了引入变化的新方法。这些方法统称为基因工程。现在能够将基因（DNA序列）从一个生物转移到另一个生物，也能在实验室合成新的基因并将其植入生物的染色体。这样就能改变基于DNA信息所表达的身体的某方面特征，使得生物具有新的特征。人类设计的变体的引入与复杂引擎的循环完全相容。从计算的角度看，人类修补匠只不过是提供了新的变化来源。

表**8.1**　自然生物系统可遗传变化的主要来源

变化分类	变化的本质
复制错误	已有核苷酸序列的变异
修复错误	已有核苷酸序列的变异
段副本	已有核苷酸序列的副本
段删除	核苷酸序列丢失
段重排	改变 DNA 分子内部序列的顺序
重组	已有 DNA 变体的重新组合
染色体的随机配对	祖父母染色体以及相应的 DNA 序列的重新混合 水平基因
转移	从其他生物补充新的基因（可以是天然的也可以是人工的）

注：列出的每种变化都对DNA有影响，也都涉及随机变化

进化算法为展示进化系统有可能用到的变化机制的谱系提供了实验室。表8.1中列出的所有自然方法都有相应的模仿算法，也有一些算法是计算机背景下独有的。单点的随机变化与重组的组合为大部分计算问题提供了足够的变体[2]。计算机实验和基因工程的最新进展都令人信服地证明了，只要技术上可行，在进化系统中人工创造新的变体在理论上没有障碍。

选择必须是天然的吗？

自然选择在生物进化中的关键作用已无需赘言。不过有趣的是达尔文之所以确信自己的思想是正确的是因为看到了动植物培育的成就：换句话说，是发现了非自然选择的力量。经过选育的狗、马和鸽子的多样性在他的《物种起源》的论证中扮演了重要角色 [3] 。很显然，就算选择不是“天然”的，复杂引擎的循环也运转得很好。研究细菌遗传学的科学家经常会培育能在毒化学环境或父代菌株不能生存的营养液中生长的菌株。培育新菌种的标准流程包含对选择的控制。一个小实验室就能培育数十亿细菌。由于突变，即使菌株全部是源自单个细胞也会有许多差异。如果每一代每个个体有一个突变，数十亿细菌的DNA序列总体上就会有数十亿不同之处，尽管许多突变是有害的，会从群体中消失。

要选育出具有新特征的菌株，诀窍是让培养基能抑制正常细菌或“野”细菌的生长，并让具有期望特性的细菌能增长得更快——只需要快一点点就够了。例如，如果想培育出对青霉素具有耐药性的菌株，青霉素加太多或太少都不行，太少了细菌都不会受到抑制，太多了所有细菌都会把杀死。青霉素的浓度要能杀死大多数细菌，但又不能全部杀死。通常，活下来的会生长，但是很慢。能生长是因为它们与大多数细菌有不同之处。但在生长和分裂过程中，它们也会突变。其中一些突变会让生长变快。生长快速的细菌很快会超过生长缓慢的细菌，从而在培养基中占据主导。然后微生物学家就会增加青霉素的浓度，杀死其中大多数细菌。同样这次的浓度也要让一小部分能缓慢生长。这些细菌的一些（随机）突变会再次使得其在高浓度青霉素中快速生长。反复执行这个过程，微生物学家很快就能培育出青霉素耐受力比最初高100倍的细菌。

这样的过程在无意中发生会给现代医学带来挑战。抗生素是20世纪最伟大的医学成就之一。这种化学药剂被广泛使用，但经常剂量不当。病人也经常不遵医嘱。结果可能无法杀死所有的入侵细菌，活下来的那些可能具有很强的耐药性。如果它们被传染给也不遵医嘱的某个人，又可能选择出更强的耐药性。如果反复以亚致死量施加抗生素，对抗生素极具耐药性的菌株就会传播开来。进化产生了，而医生则失去了抵抗感染的一个重要手段。

这个现象不限于细菌：HIV（艾滋病病毒）会对抗病毒药物产生耐

药性，疟原虫对抗疟疾药物，蚊子对DDT也是如此。突变加选择的重复循环是改变已有基因的有效工具，但在创造全新的基因时并不是很有效。要做到这一点最有效的方法是从其他地方获得新基因。要在有抗生素的环境生存，最好的办法是基因的表达能摧毁这种化学因子。大多数抗生素是从土壤真菌中分离出来的，而土壤细菌与土壤真菌已经共存了很长时间。结果许多土壤细菌进化出了能编码蛋白质摧毁特定抗生素的基因。细菌可能对某些抗生素很敏感，而对它们的耐药基因所针对的药物又极具耐受性。大多数细菌都有与其他细菌交换基因的机制。这种交换使得细菌能完整获得有用的基因。在适当的条件下，基因转移机制能让基因在菌群中迅速传播。通常，耐抗生素基因被一起封装在名为质粒的单独的DNA片段中，这些质粒能在细菌之间传递。因此，大量使用抗生素的地方，例如医院，通常会有很多耐抗生素的细菌。这是选择所引发的自然的进化结果，也是为什么在医院被感染比在家里危险得多。

现代农业是基于自然选择再进行人类选择。在大约10000年前的农业革命中，人类学会了增强进化过程。在农业刚开始时，农民可能会吃掉好的种子，把差的留到来年播种。然而，肯定会有一些人发现更好的做法，将今年最好的玉米种子选出来留到来年播种，以及让最好的动物繁殖，宰杀那些不那么好的。不这样做的人很快会发现他们的玉米产量比其他人差，从而选育最好的就成为了农业的标准做法。通过一年又一年一代又一代施加人工选择，玉米和家畜逐渐变得远超它们的野生祖先。

对谷类作物的选育不仅使得产量远超过它们的野生祖先，同时也表现出了许多更适于人类消费的修正。其中包括统一发芽（这样才能赶上播种），谷物的秸秆既强壮又脆（这样到收割时谷物还能留在秸秆上，但又易于用机械分离），不那么坚硬的种壳（便于脱壳和碾磨），植株形态更加挺拔（便于收割）。平行的变化同样发生在家畜身上。家畜通常会变得比野生种群更大，更多肉，更温顺和更易于繁殖。

在近几个世纪，遗传学知识和现代统计手段的应用使得传统选育进一步改善。结果，美国的玉米亩产在100年里翻了5倍。1905年亩产是760千克，现在是3800千克。部分是因为技术的进步，大部分则是由于遗传学的进步。其他主要玉米品种也有类似的故事。“绿色革命”经常被认为是全世界人口迅速增长而饥荒却在减少的主要原因。现在奶牛年产奶20000磅很平常，而在100年前则是闻所未闻。这些奶牛是复杂的选育

策略的结果，建立在大量数据的基础之上，用到的信息包括数百万头奶牛的产奶量、营养、抗病性，等等。这些信息使得能导致期望特性的基因可被追溯许多代，并识别出最佳的动物组合进行繁育。背后的原理与传统农业是一样的：让具有最佳基因的动物一起繁育，只是现代技术要高效得多。

现代动植物选育使用的是与第5章一样的复杂引擎，只是选择不是天然的。动植物选育者控制了循环的选择部分。经验表明几乎动植物的任何特性都能通过对选择的一致控制进行修改。除非特性没有表现出可选择的变化或者完全缺乏所期望的特性时才有例外，因为没有东西可以修改。

现代生物技术可以在生物的DNA中引入全新的基因。用这种方法可以产生出能被选择的新特性。自然的进化循环依赖于随机的变化产生新的变体；基因工程则能够引入预先确定的变化。结果是大大加快了变化的速度。产生所期望的特性不再需要成百上千代，而只需要很少的几代就行。一些例子已经被商业化，包括蛾幼虫无法消化的玉米，耐受草甘膦（一种能杀死大部分植物的除草剂）的大豆，以及维生素A含量很高的大米。从某种意义上来说，这种技术只是模仿了细菌交换基因以获得抗生素耐受性的策略。

计算机中运行的遗传算法将对选择的控制推到了极限。所有在对生物的人工选育的例子中，人类选择都只是自然选择的补充。当奶牛死于疾病时，是自然选择在起作用；当公牛被农民选出来配种时，是人工选择在起作用。当计算机学家写进化算法时，选择的规则完全是指定的；当程序运行时自然没有参与，但进化仍然发生了。

循环有哪方面不能控制吗？

复杂引擎的循环包含几个方面（见图5.2）。有可被复制并用一致标准进行评估的信息体。在大多数系统中，信息编码为可以用来创造某物的指令。同时还有复制机制，可以产生多个复制体，并且其中一些可以发生变化，然后还要有选择，不是所有的都能继续复制。

选择的过程不能随机，必须具有一致性。这并不是特别复杂的要求；在许多场景中都应用了这一策略。

由于复杂引擎本质上就是计算，因此在编码和复制信息的任何系统中都能引入逻辑。这也意味着只要知道编码、复制和选择的本质，就有可能进行干预，引入人为的变化、新的复制方法，或者替换选择规则。一旦信息编码为指令形式，就总是有可能利用这个循环得到所期望的新结果。换句话说，控制是有可能的。

在人类对生物王国的干预中，生物技术完全控制了这个循环的所有方面。从基因组（DNA）开始。除了控制选择和变化机制，最终还会实现整个基因组都根据要求定制。你可以想象基因工程师将整个DNA序列输入计算机，直接将细胞完整的DNA合成出来。然后这个合成的分子被植入一个原来的DNA已被移除的细胞中，然后在实验室或野外进行培育。所有关键技术，包括合成DNA序列的能力以及将这种分子植入细胞的能力都已经具备了。目前在实验室中已经人工创造出了病毒和小细菌的基因组，并植入了细胞中。

可以确定的是，如果这样的生物被创造出来了，如果可以繁殖，还是不能避开自然选择。还可以确定的是，除非我们的生理学和生态学知识远远超过了目前的水平，否则将人工创造出来的生物放到野外肯定很快会灭绝。一旦人为选择被除去，自然规则重新建立，野生的表亲会很快证明它们在旧规则下的优势。

在最后的几章里我们将探讨人类学习和社会变迁背后的复杂引擎。如果这些也是受复杂引擎驱动，社会工程和改善人类思维的技术的主要障碍就是缺乏对信息如何在人脑中编码的清晰认识。如果能有更深的认识，人类将很有可能找到方法在相关的信息体中植入特别设计的变化，或者发现改变选择规则的方法，从而改进大脑功能和/或社会组织。但在探讨这些问题之前，我们需要先探讨另一个越来越重要的现代科学概念——复杂系统。

第9章　复杂系统

复杂系统是什么？

复杂系统没有广泛认可的数学定义。计算机、雷暴、人脑和公司都是复杂系统。它们都是由许多以特定方式互动的部分组成。这些互动的结果通常很难预料。我的一个目标是解释复杂事物的存在（而事物与系统之间没有明显的区分），因此有必要深入了解一下这类系统。显然，系统的组分越多，表现出复杂行为的可能性也越大。这个领域一个令人吃惊的发现是，复杂的组分对于复杂行为并不是必需的。当系统有许多互动的组分，即便组分和互动规则很简单，有时候也足以产生出复杂的结构和行为。第2章的NKS计算系统和第3章的沙堆就是这样的例子。

描述复杂系统的一个重要手段是网络。只要系统可以被刻画为组分之间的互动，系统如果足够简单的话，就可以用网络描绘。数学家将网络称为图，个体组分称为节点，节点之间的互动则称为边，图9.1给出了一个有9个节点和10条边的简单的图。

图的数学概念很广义，只要是有互动的组分就可以用图来描述。图可以是任意维度（大于0）。网络的一个特性是，如果互动（图的边）是非线性的，则通常会有包含多个稳态的复杂行为。

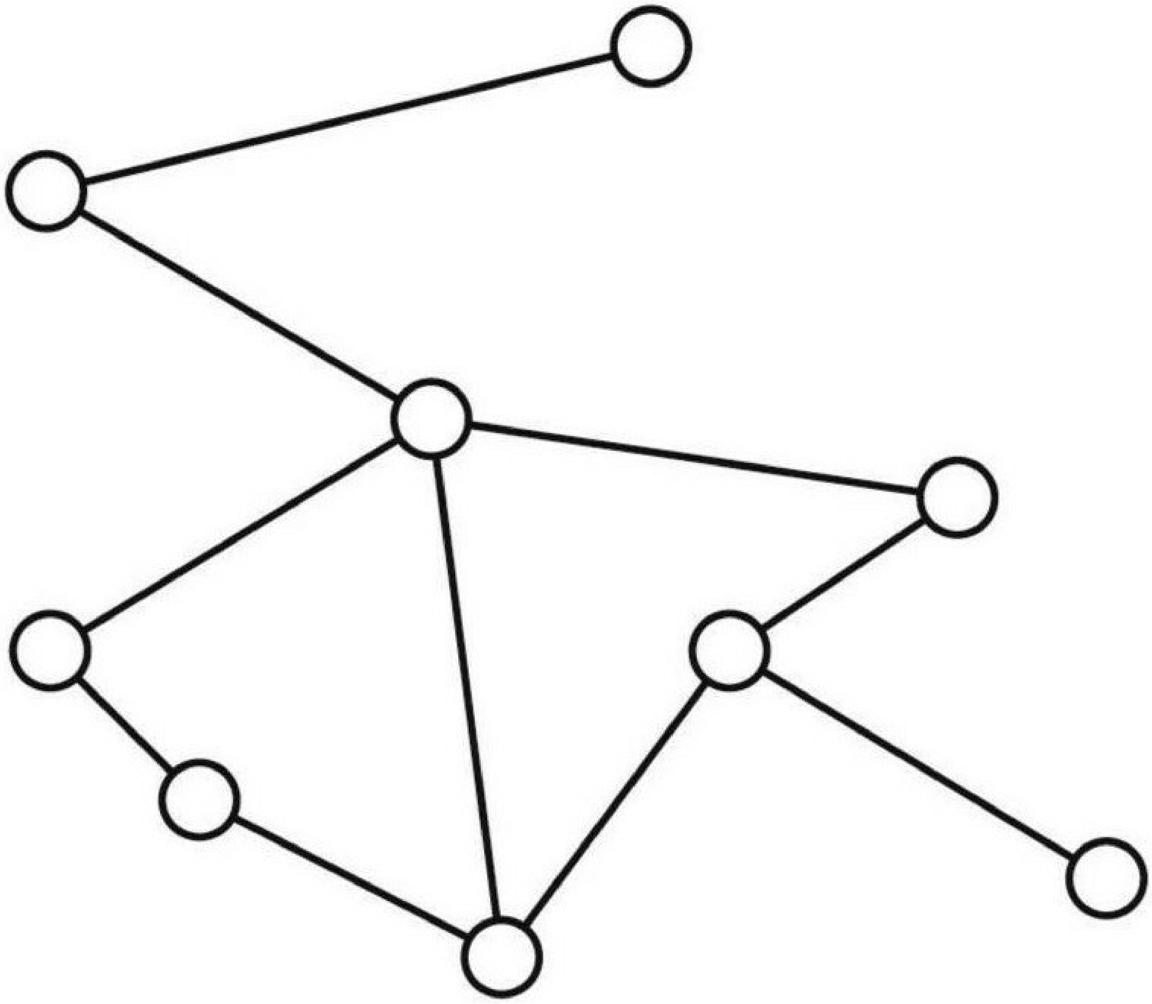


图9.1　9个节点和10条边组成的图

斯图尔特·考夫曼研究了特定类型网络的行为，尤其是布尔网络，其节点根据布尔逻辑进行互动。考夫曼向大众介绍了这种大量简单组分根据简单规则互动导致的自组织特性[1]。考夫曼用“混沌边缘”一词描述其中的一项原理。大规模网络的特点是很大部分参数空间都会导致混沌行为，还有一大部分被不变的行为占据；但是在不变的和混沌的区域之间，会出现复杂的行为——甚至可以说是具有创造性。考夫曼认为进化系统——尤其是生命——就位于这样的动力学参数空间区域。他称之为“免费的秩序”。

也许是因为考夫曼的大部分研究是基于简单系统，他不太重视具有复杂组分的系统表现出的丰富行为。在自然界有很多系统是由简单组分通过简单互动自行产生秩序并表现出丰富的行为；但是在生命和人类活

动领域，我们看到由复杂组分组成的系统具有的惊人潜力。当网络节点复杂到需要指令才能形成，就为全新的行为世界带来了无限可能。为了揭示这一点，我们将探讨6种网络，其中最后一种为生命提供了定义。在每个例子中，都是预先设定好网络的组分（节点或边），使得特定的网络特性成为可能。互动的规则就是化学和物理定律。

为什么电气设备也是网络？

我们考虑的第一个网络是电路。电路很容易用图表示；组分（节点）是电阻、电容、开关等元器件；组分之间的互动（边）是通过导线的电流。

开关灯

一个很简单的电路是电池与开关和灯泡串联（图9.2），合上开关就会通电。这会使得灯泡发光发热。切断开关则会截断电流，灯熄灭。这个网络有3个节点（电池、开关和灯泡）、3条边（连接节点的导线）、2个稳态（开和关）。

图9.3展示了由相同元件组成的3种不同电路。左边的电路在开关合上后没有变化。中间的电路在开关合上后只会导致电池短路，而右边的无论开关是否合上电池都会短路。这3个电路都没有灯泡能点亮的状态。总的来说，只有3个元件串联，才能在合上开关时让灯泡发光。有很多图都可以由这同样的3个节点组成，却并不具有这个特性。

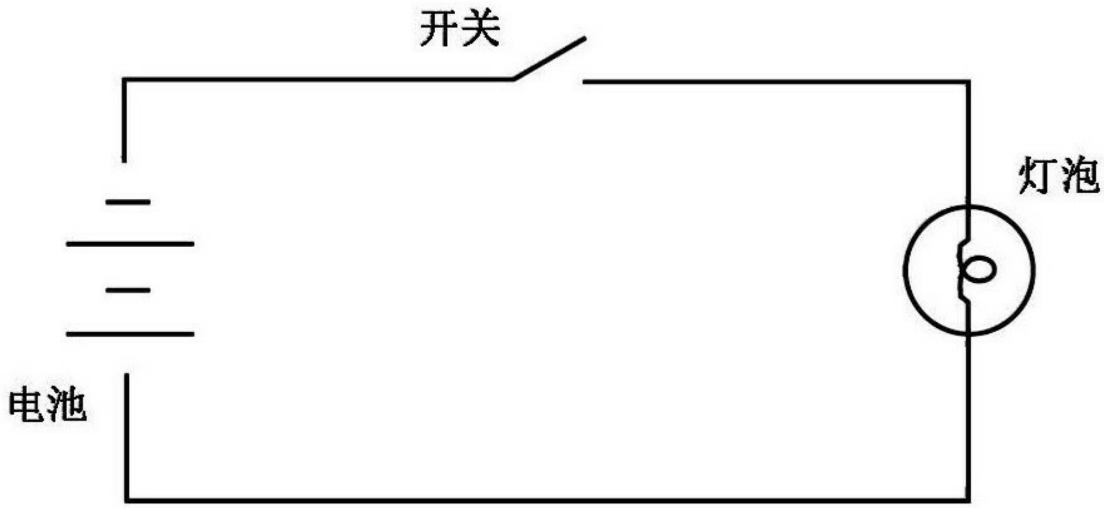


图9.2　3个元件串联而成的简单电路

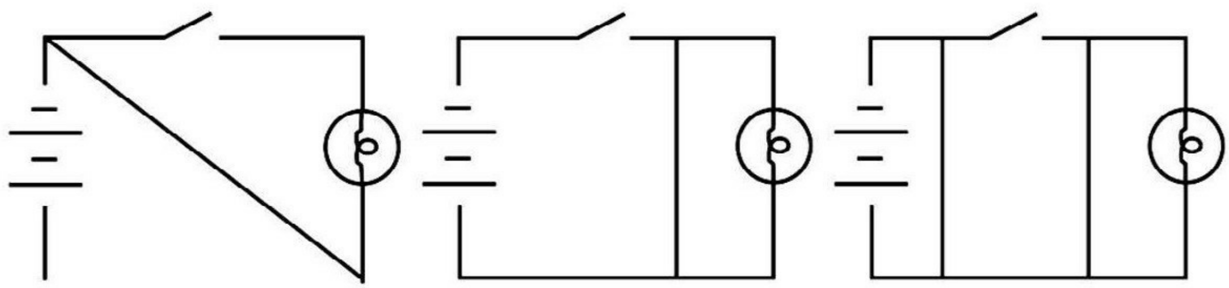


图9.3　图9.2中同样的3个节点组成的3个电路并没有表现出值得注意的行为

晶体管收音机

现在考虑如图9.4所示由14个节点组成的电路。这是一个能将AM无线电波转化成声音的收音机。这个网络由以下部分组成：天线、线圈、二极管、三极管、可调电容、3个固定电容、电池、耳机、3个电阻以及以特定方式将这些元器件连起来的导线。电流从电池出来流经三极管和耳机。

如果流经耳机的电流迅速变化，就会产生声音。可变电容和线圈让操作者可以将电路调谐到天线获取的某个广播信号频率。三极管电路放大天线获得的微弱电振荡，从而可以在耳机中产生听得到的声音。

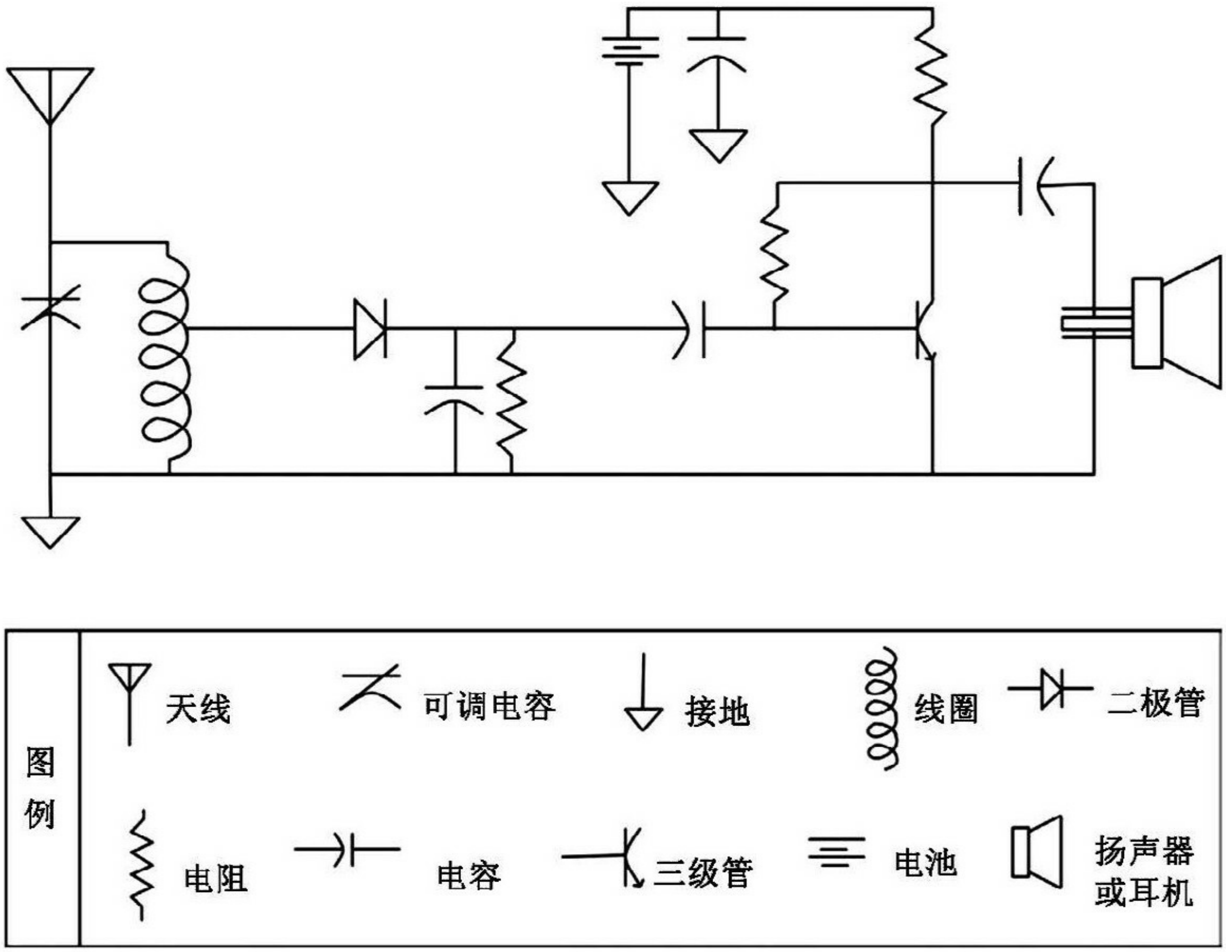


图9.4　简单的晶体管收音机电路图。参见：
http：//www.techlib.com/electronics/crystal.html#Crystal%20 Radio

这个网络的输出很稳定，随着天线捕捉到的信号变化。这个网络有大量稳态，可以通过可调电容设定。有许多方式可以将这14个元件连接到一起，但大部分都不具有收音机的功能。用其他元件也能组成收音机，但不具有收音机功能的电路数量举不胜举，而具有AM收音机功能的电路数量则相对很少。由于在所有可能的电路中，收音机与非收音机很不成比例，因此我们不能期望收音机能自己出现，除非有某种机制说明了元器件以及它们的连接。

与收音机相比，能间歇发光的电路则可以在自然界中遇到。闪电的电路图很类似图9.2，只是云层中的电荷替代了电池，穿过空气的通路则替代了开关、电线和灯泡。当空气被电流加热，就会变得更容易导电（开关），一旦空气被充分加热，就会变得白热（发光）。在自然界中发现的电路与人类创造的电路之间的区别是复杂性。据我们所知，所有

复杂的电路都是人类创造的，其行为要比简单的振荡或光的亮灭复杂得多。所有人类创造的电气设备的生产也都需要指令。

我们再来看几个无需人类干预也能表现出复杂行为的目的性网络。

生物化学网络是什么？

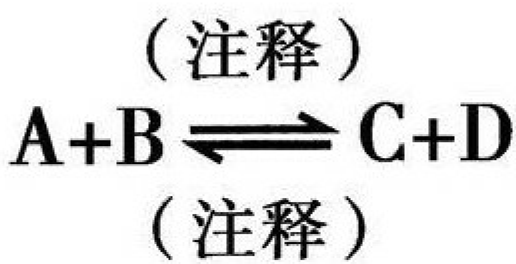
化学反应通常表示为下面这种方程式：A+2B⇒C

C+D⇒E+F

A+E+F⇒G+H

每个字母表示一种化学物质或分子类型，例如水（H_2O）、氨（NH3）、苯（C6H6），等等。化学方程式可以告诉我们一个A分子可以与两个B分子结合形成C分子，但没有告诉我们这个反应到底会不会发生和有多快。要预测到底会发生什么，化学家还需要更多信息。一个重要的信息是每种反应物的浓度。一个基本的化学原理是，如果反应A+B⇒C能发生，则逆反应C⇒B+A也能发生。实际的方向取决于3种反应物的相对浓度以及背后的能量关系。有了这些信息，化学家就能知道化学反应的方向，但还是不知道反应的速度有多快。

要确定反应的速度更加复杂。它取决于温度、一个名为活化能的特性、所有反应物的浓度以及参与反应的每种化学物质的几何特征。甚至只间接参与反应的分子也对速度有影响。通常处理所有这些复杂因素的办法是将化学反应表示为以下形式：



其中向右的箭头的注释是关于正向反应的重要信息，向左的箭头的注释则是关于逆向反应的重要信息。最常见的注释是“反应速率常数”，综合几个参数得到的一个数字，但注释也可以提供其他信息，例如是否有催化剂（反应加速剂）、特殊溶剂等对反应特性有重要影响。

化学反应方程式是1维图。当数个反应一起发生，并且共用一些化学物质时，系统行为会变得很复杂。用2维或3维网络表示可以帮助我们理解发生的事情。图9.5a给出了一个有8种化学物质参与的简单反应，其中A、C、E和G各参与两种反应。图9.5　b用2维网络表示了同样的反应。双向箭头表示每种反应的方向都是由反应物和产物的浓度决定。

通过图9.5　b可以发现在反应物中加入更多的F最初会增加G的浓度，这又会使得A和E增加，递次推进。因此F的浓度增加通过网络最终会导致B、D和H的浓度增加。如果增加F并减少B，则会形成从F→G→（A和E）→C并最终到B的物质流。在这个过程中，每种化学物质的浓度都会变化。

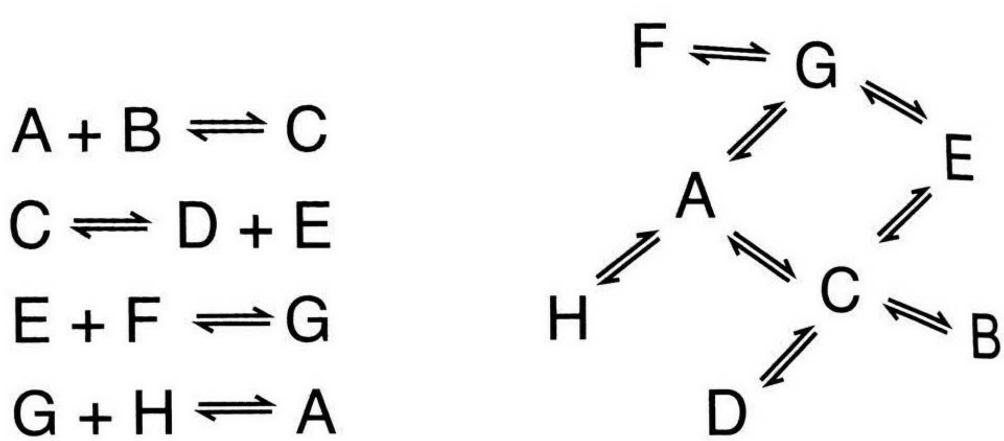


图9.5　a.有8种化学物质参与的4种化学反应b.用2维图表示的a中的4种反应的8个关系

这些变化的细节牵涉整个网络，仅从图9.5a很难看出来。当多种化学物质以多种方式反应时，有些物质的浓度可能很稳定，其他物质剧烈变化时也不怎么变，有些物质的浓度则可能表现出振荡，甚至更复杂的行为。当系统中存在多种反应时，表示成网络有助于将物质成信息流可视化。

当节点和边的数量太多时，描绘的化学网络会变得很混乱；但即使是这样，通过有选择地略掉一些边，网络视角还是能帮助人们从整体上把握过于复杂的化学网络的逻辑。细胞代谢网络就是一个很好的例子。活细胞中有数千种不断反应的化学物质。这个网络在一定程度上决定了细胞的性质和功能。化学物质浓度的涨落取决于最近吃了什么以及细胞生命的各种需求。总体上，网络为基本的细胞活动提供稳定的化学供给。所有细胞化学网络的一个共同的特性是所有化学反应都可以被蛋白质酶催化和调控。每种酶针对特定的化学反应，控制反应的速度（但不是方向）。因此，在生物化学网络的每条边（反应箭头）上标注控制这种反应的酶会很有用。

由于酶的活性通常受网络中其他一些特定的化合物的浓度调控，因此在绘制细胞生物化学网络时可以添加另一种边——“逻辑边”。逻辑边表示控制、信息流而不是物质流。这些边的信息“决定”了物质边上的物质流的增加或减少。调控是通过改变特定酶的物理性质实现。图9.6展示了一个庞大的生物化学网络的一小部分。这个子网络包括12种化学物质，还有2条逻辑边。为了减少边的复杂程度，三磷酸腺苷（ATP）、二磷酸腺苷（ADP）、二氧化碳（CO2）和无机磷酸盐（P_i，HPO_4^-和$H_2PO_4^-$的混合物）出现了多次。完整的细胞生物化学网络包含的化学物质超过2000种。

维持生命的生物化学网络缺少不了酶。如果没有酶，在生命所处的温度范围内，细胞的大部分化学反应都不会发生，而这些酶很少有不受调控的。反应速度的调控对于生命很重要。将这2000种细胞化合物倒进烧杯，在室温下什么也不会发生。加热混合物，结果将是一团混乱，而不是适时适地提供的细胞生长和复制所必需的经过精细调制的化合物流。

有特定的蛋白质才能进行组织和控制。在第4章我们看到DNA分子中的核苷酸序列是如何决定蛋白质的结构和功能。DNA编码的信息决定了在细胞中会表达什么蛋白质，以及在特定的时间产生多少蛋白质。其中一些蛋白质反过来又决定了生物化学网络会发生什么反应以及速度如何。最后的结果就是能维持细胞特有的稳定化学环境的系统。

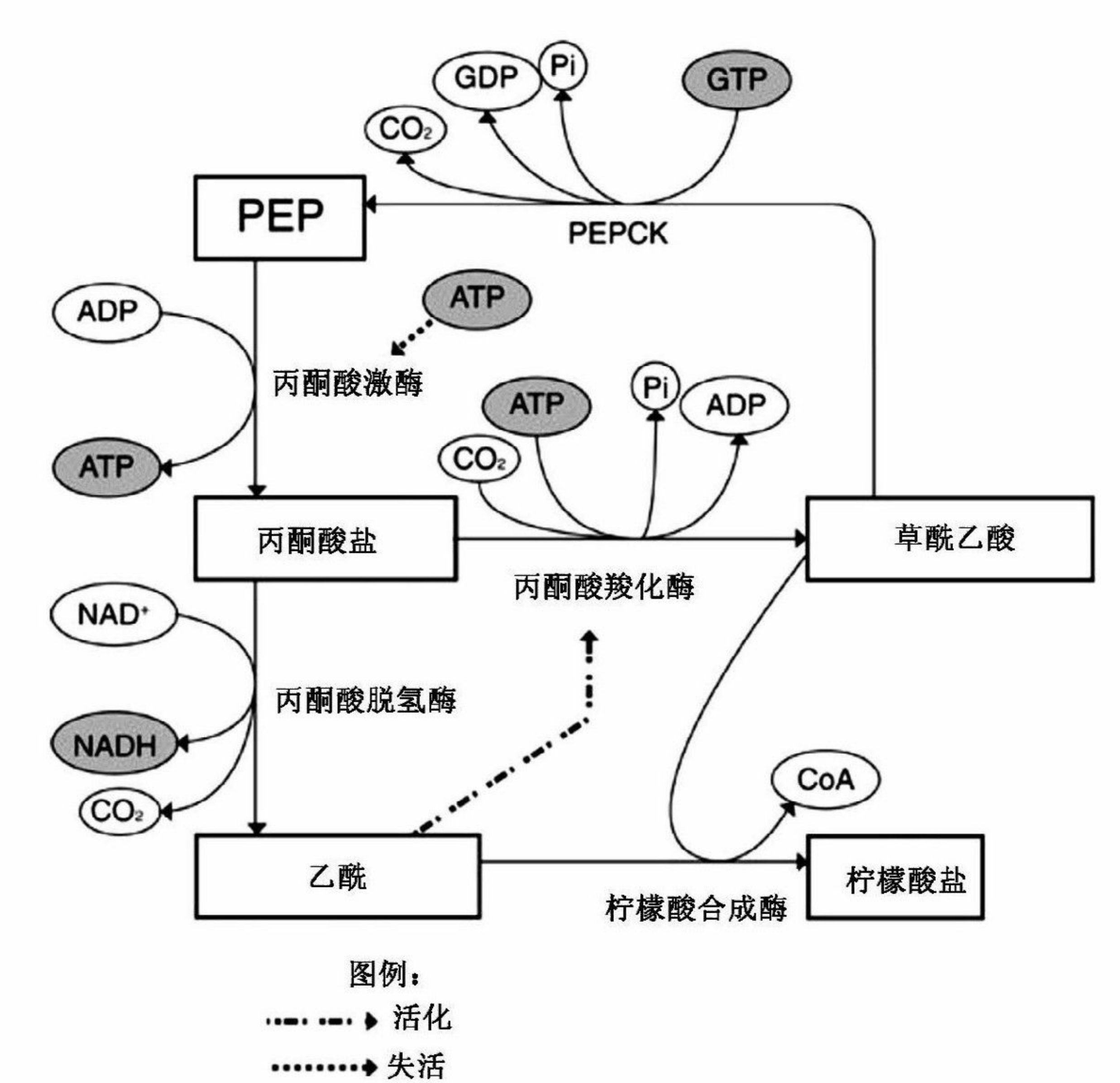


图9.6　人类肝细胞的生物化学网络的一小部分。实线是物质边（化学反应），虚线是逻辑（调控）边，节点是化合物，边标注的是酶。缩写：PEP，磷酸烯醇丙酮酸；PEPCK，磷酸羧；CoA，辅酶A；ATP，三磷酸腺苷；ADT，二磷酸腺苷；GTP，三磷酸鸟苷；GDP，二磷酸鸟苷；NAD^+，烟酰胺腺嘌呤二核苷酸（氧化形式）；NADH，烟酰胺腺嘌呤二核苷酸（还原形式）；P_i，无机磷酸盐；CO_2，二氧化碳

基因网络是什么，又是如何运作的？

你的细胞中的DNA编码了生成大约22000种不同蛋白质所需的信息。其中大约有10%的多肽是酶，其余的则有其他的用途。大肠杆菌的

DNA编码了4322种蛋白质。有些生物的基因比人类还多，有些则比大肠杆菌还少，但基本都包括了在地球上产生和维持生命所需的各种基因（编码蛋白质的指令）。

即使最粗心的人也会意识到，无论从哪方面来看，人类都要比细菌复杂不止5倍。大肠杆菌只有一个细胞，而一个成年人大约有100亿（10^{13}）个细胞。人类细胞包含各种复杂的亚细胞结构，包括染色体、细胞核、线粒体、复杂的囊泡输运系统以及由上百种相互作用的蛋白质组成的细胞骨架。细菌细胞则没有这些结构。人体有大约200种细胞类型，组织成各种结构、器官，以及在身体各部位之间进行通信和协调的神经和内分泌系统，与入侵身体的微生物作战的免疫系统，可以迅速移动的骨骼肌肉系统，还有管理身体所需的各种物质的消化、呼吸、循环和排泄系统。人类大脑被认为是宇宙中最复杂的事物（更不要说以大脑为组成部分的人体）。大肠杆菌这些都没有。现代基因组学面临的一个最大难题就是理解只比大肠杆菌多5倍的基因如何表达出人体这样的复杂性。

答案目前还不完全清楚，不过可能涉及几个方面。首先，多细胞动植物通过一种名为选择性剪接的过程可以生成大部分蛋白质的多个修饰版本。在这个过程中，在合成蛋白质之前根据基因生成了信使RNA的多个版本。目前还不清楚选择性修饰对于多细胞复杂性的产生有多重要，不过据估计这个过程使得人体产生的多肽的种类增加了100000种。

要理解细菌和人类的差别，最重要的可能是基因调控的相对复杂性。包括大肠杆菌在内，所有细胞中的基因都有一个重要特征，就是基因受控在哪种环境下表达。在任何特定的人类细胞中，在任何特定的时刻，22000种基因中只有1/3到一半会表达，并且如果两种不同的细胞表达了相同的蛋白质，蛋白质的数量也有可能不一样。随着时间推移，蛋白质表达的不同调控，能在由数千节点组成的网络中产生具有不同特性的稳态。

所有动植物的独特复杂性都可以回溯到单细胞时期受精卵或结合子。初始细胞完全不同于成年人体中发现的数百种不同的细胞类型（除了卵巢中成熟的卵细胞）。从受精卵开始形成成年人体的过程中，同时发生了4种活动：细胞通过分裂大量增殖，然后通过细胞分化过程不断改变细胞的物理特性，细胞通过协同移动和生存或死亡在预定位置形成

组织和器官，生物从食物分子中获取维持生物化学网络所需的新物质。

通过这4种活动，新生成的细胞类型所建立的结构的互动逐渐产生出身体的复杂性。细胞类型与结构的每次互动又进一步导致结构的细化和细胞类型的多样化，在之前建立的复杂性的基础上建立起更大的复杂性。

所有身体层面的变化背后是具有特定性质的单个细胞的行为。这些性质反过来又是由细胞中表达的蛋白质的数量和类型决定。蛋白质决定生物化学网络的细节，同时也参与细胞中表达的各种结构。单个细胞性质和行为的变化决定了整个身体的形态和能力，因此最终是细胞中22000种基因的不同表达和活动决定了我们是谁。基因组分析表明，用于基因表达调控的人类DNA的数量是直接编码蛋白质的DNA数量的3到5倍。而大肠杆菌只有不到10%的DNA用于基因表达调控。

很显然，要（从分子层面上）认识我们自己，就必须理解基因表达的调控。图9.7是一个基因调控的示意图。调控是通过控制RNA聚合酶实现，RNA聚合酶在细胞中的作用是生成部分DNA的RNA复制体。第4章曾说过，核糖体根据信使RNA分子合成蛋白质。在蛋白质合成过程中，信使RNA分子上核苷酸的顺序决定了形成蛋白质分子的氨基酸的顺序。每个信使RNA可以合成数百个同样的蛋白质分子。由于信使RNA分子会退化，新的信使RNA分子必须不断形成，才能不断生成特定基因对应的蛋白质。因此，蛋白质表达的调控主要就在于信使RNA的生成。

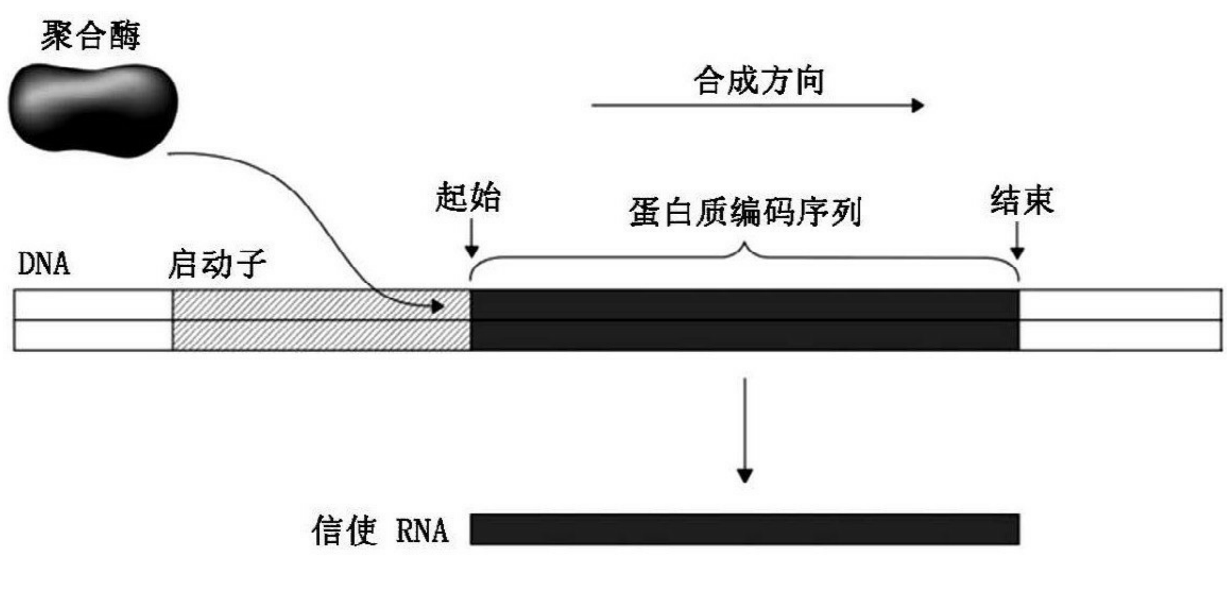


图9.7　与蛋白质编码序列相邻的控制序列的典型基因结构，被称为启动子。信使RNA是DNA蛋白质编码序列的复制体

图9.7中双横线代表DNA分子中的一段。RNA聚合酶不断在DNA周围游荡寻找基因。“起始”和“结束”序列标识RNA聚合酶开始和停止复制DNA的位置。复制总是如图所示从左往右进行。一段称为启动子的DNA序列决定RNA聚合酶是否能成功找到旁边的起始位置。RNA聚合酶的启动有多频繁决定了生成的RNA分子的数量，从而又决定了生成的蛋白质分子的数量。基因表达的控制是通过控制RNA聚合酶是否能到达起始位置实现的。

细胞中与DNA启动子序列结合的蛋白质被称为转录因子。不同的启动子区表现出不同的转录因子结合区，因此细胞中的各个启动子都有独特的转录因子组合。转录因子可以阻止（抑制）或增强（激活）RNA聚合酶的活动。启动子的各种抑制子和激活子的平衡决定了RNA聚合酶在相应的基因的活动。由于抑制子和激活子本身也是蛋白质，因此细胞中表达的抑制子和激活子的数量又是由其他抑制子和激活子决定。这是一个无穷递推系统。所有转录因子的数量都是由所有转录因子整体的浓度决定。因此，基因调控组成了网络，与生物化学网络很类似，只是节点是基因而不是生物化学物质。

这个网络的逻辑结构决定了能生成多少信使RNA。网络响应细胞环境的变化，调控合成哪种蛋白质，合成多少。细菌（例如大肠杆菌）和高等生物（例如人类）的一个重要差别是影响每种基因表达的转录因子的数量。对于大肠杆菌，通常是2种。对于人类基因，数量通常超过20种。由于多细胞生物（例如人类）的基因网络具有更多调控边，网络的复杂度也就比细菌（例如大肠杆菌）要复杂得多。也许就是基因网络的复杂性的差别使得人体比细菌要复杂得多。图9.8给出了一个简单的基因网络的例子[2]。

同电路和生物化学网络一样，可能的基因调控网络的数量是无法穷尽的。最近的基因组测序表明老鼠和人类的基因（节点）在很大程度上是一样的。主要差别在于基因调控网络的结构。网络的启动子序列和与序列交互的转录因子的微妙结构差别决定了网络的性质。由于转录因子和所结合的DNA序列都编码在DNA中，因此基因调控网络在整体上都是由DNA分子中核苷酸序列写成的指令决定。如果没有指令，就不会有网络，也不会有生命。

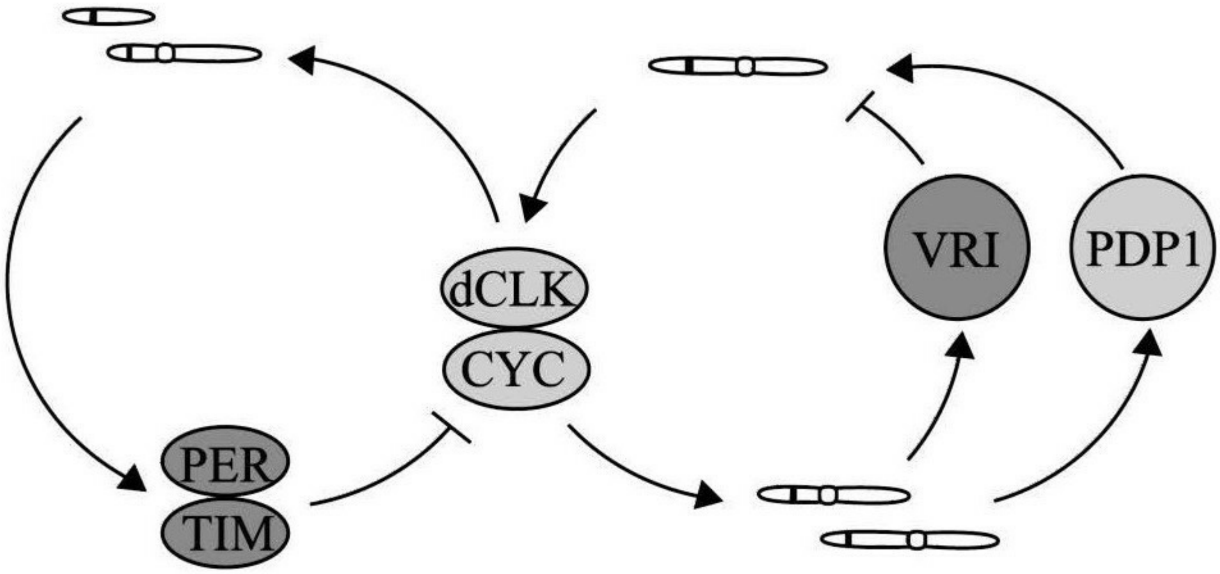


图9.8　生物钟。果蝇的生理节奏（24小时）循环是基于椭圆表示的6种蛋白质的循环表达。CYC和dCLK相互结合，并激活染色体（细长结构）中编码蛋白质PER、TIM、VRI和PDP1的基因的表达。PER和TIM形成二元复合物，抑制CYC/dCLK的复合。VRI抑制dCLK和CYC的合成，而PDP 1则刺激dCLK和CYC的合成。这个网络的蛋白质浓度以24小时为周期振荡（参见注释2）

蛋白质组是什么？

要完整认识生命还需要第3个网络，蛋白质的相互作用，有时候也称蛋白质相互作用组，或简称蛋白质组。在分子层面上，细胞中发生的一切几乎都是因为蛋白质与特定的某物结合。这个简单的思想构成了现代生物学观念的基础。酶特异性结合酶作用分子，基因调控源自转录因子与特定的DNA序列结合，RNA需要特定的结合蛋白才能发挥功能，建构细胞结构时是特定的蛋白质与蛋白质结合。

大部分蛋白质都与其他蛋白质结合。这种相互作用网络就是蛋白质组。一些蛋白质只与少数蛋白质交互，有些则与数百种交互。由于酶的活动不仅受小分子的交互影响，也受其他蛋白质的交互影响，因此酶就成为了蛋白质相互作用组和代谢网络的交叉点；而由于转录因子除了DNA还与其他蛋白质结合，因此基因调控网络与蛋白质组也相互交叉。因此，最好将这3个网络视为更大的超级网络的一个侧面。分析这个超级网络的结构和动力学对于认识细胞很重要；事实上，这就是细胞。由于这个超级网络是由蛋白质界定，而蛋白质又取决于DNA序列，因此这

个网络的最终特性取决于DNA中编码的指令。DNA也是这个超级网络的组成部分，因为DNA是由蛋白质酶复制和修复，而且DNA还为基因调控网络提供了节点。

生命是什么？

据我们所知，所有生命都是基于细胞。生命与细胞的关系是如此紧密，可以说地球生命就是由这个我们称为细胞的分子活动的小包囊界定的。一些生物由单细胞组成，另一些包括我们自己则是由大量合作的细胞组成。由生物化学、基因、蛋白质3部分组成的超级网络界定了细胞，决定了细胞如何对环境做出反应，如何随时间变化，以及如何修复自身，但没有解释细胞是如何开始的。要从分子层面理解生命还需要一个要素，就是已存在的结构。在第4章我们看到特殊设计的蛋白质是如何自组装成噬菌体。细胞通过蛋白质自组装机制生成大量结构，包括多酶复合物、核糖体和大部分细胞骨架结构，以及各种细丝和用于运动的分子马达。同噬菌体类似，如果合适的蛋白质（有时候还需要RNA）在适当的条件下在试管中混合，这些结构就能自组装。然而，有一些细胞结构无法自组装。其中包括DNA以及基于生物膜的大部分结构，例如线粒体和叶绿体。细胞骨架结构的某些方面可能也属于这一类。

所有细胞都被生物膜包裹，复杂细胞还会被生物膜分隔成多个腔室。膜的合成可以被视为是生物化学网络的一部分，在实验室条件下可以人工合成类膜结构。然而，在细胞中，新的细胞膜总是通过向已存在的膜添加新物质而产生。

生物膜主要由两种成分组成，磷脂和蛋白质。磷脂分子的一端具有亲水性，另一端具有疏水性。有很多分子具有这种特性，包括肥皂，其中大部分与水混合会形成拉伸的球，称为胶束，每个分子疏水端朝内，亲水端朝外，与结构体外面的水分子接触。磷脂有时候也会形成胶束，但大部分时候会形成名为磷脂双分子层的2维平面。图9.9a给出了磷脂双分子层的示意图。图中球和绳状的图形表示磷脂分子，球（或“头”）是亲水端，两条弯曲的线（或“尾”）表示疏水端。双层结构形成了薄薄的一层一直延伸。在磷脂双分子层中，头与水接触，尾与尾接触，从而避免了与水接触。

一般来说，磷脂双分子层无法被水和亲水分子渗透，但可以被疏水分子渗透。由于水中的疏水分子不容易接触到浸入水中的磷脂双分子层，这样双分子层就为大部分化合物的自由移动建立了一道屏障。

所有生物膜都有嵌在磷脂双分子层上的蛋白质。这些蛋白质修饰双分子层的物理特性，可以选择性通过特定的小分子和跨膜传送环境信息。图9.9 b中类似土豆的结构表示的就是膜蛋白质。实际的蛋白质结构要复杂得多，就像图4.3和图7.3中那样。

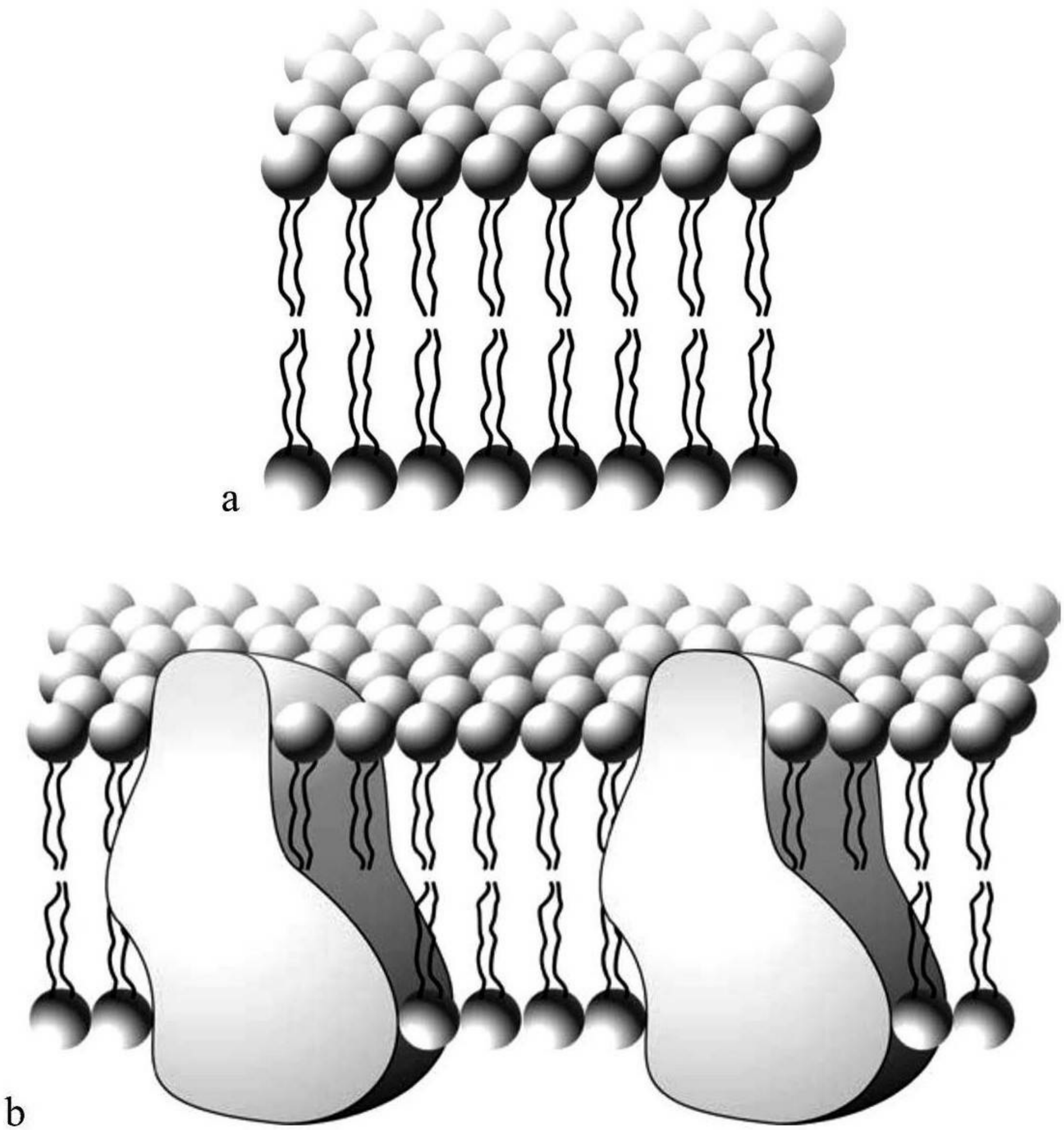


图9.9

a.磷脂双分子层

b.磷脂双分子层上的蛋白质

图9.9展示的双分子层结构是平面结构，边露在外面。这并不是稳定的状态，自然界中很少会出现这样的结构。在人工合成时，双层结构会卷曲，边会合拢形成球或囊泡，水被包围在中间。这种形态避免了边

暴露在外面，将内部充满水的腔室与外部的水溶液隔离开来。细胞就是一个大囊泡，细胞质则是围在内部的水溶液。许多细胞的内部还有大量更小的囊泡。膜决定了什么可以进来什么可以出去。所有细胞核囊泡都维持着一个与外界不同的内部化学环境。调控是通过膜蛋白质以及它们与膜内部的三元超级网络的互动实现的。

膜蛋白质的活动由它们的结构、与其他蛋白质的互动以及代谢网络成分分子的浓度决定。由于超级网络的各部分都是来自网络各部分之间的互动，从某种意义上可以说超级网络是自决的；但要做到这一点就必须将其装入由其自身构造和调控的生物膜。这个膜使得分子无法游走，因而可以创造出独特的内部化学环境。通过合成磷脂和适当的蛋白质，并将其插入已存在的膜，细胞就能生成新的膜，但如果没有预先存在的膜，细胞似乎并没有能力产生出膜。

这个只能在已存在的结构上生成的特性，DNA表现得更明显。DNA分子总是通过复制已存在的DNA分子进行合成。图9.10描述了这个过程。细胞中的DNA是双重结构，由两条很长的分子通过氢键微弱的化学吸引力结合到一起。两条长绳相互交缠形成双螺旋结构。双螺旋上的两条长绳不是独立结构，如果一条上是碱基A（腺嘌呤），另一条上则对应为T（胸腺嘧啶）。这是因为A和T能准确匹配，在中间形成氢键。同样，如果一条上是G（鸟嘌呤），另一条上则是C（胞嘧啶），G和C也能准确匹配到一起形成氢键。正是A和T以及G和C之间的氢键将两条长绳结合到一起形成双螺旋。如果它们不匹配，就无法结合。

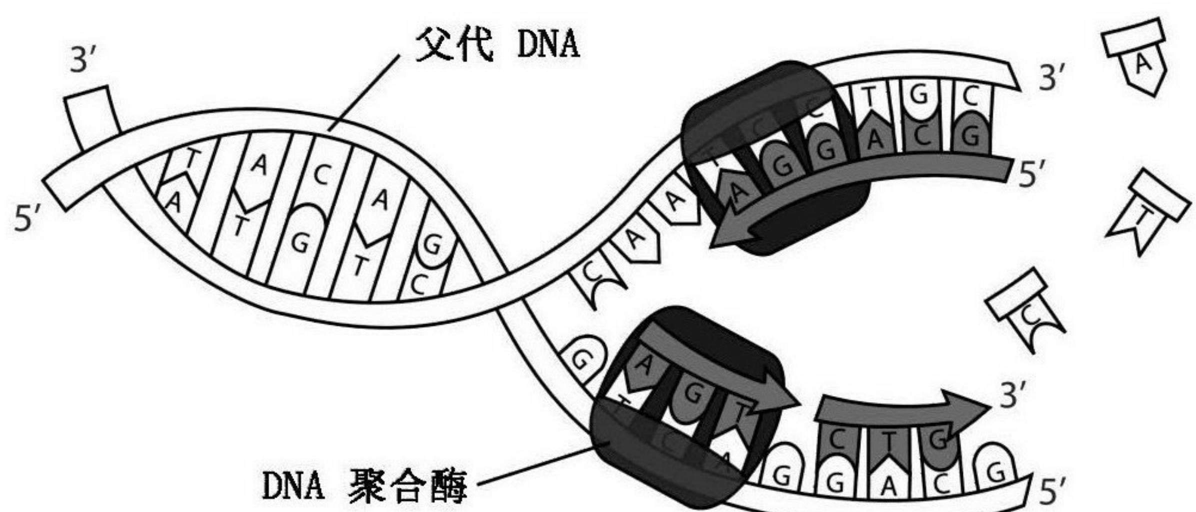


图9.10　DNA的复制。左边是原来的（父代）长绳，具有双螺旋结构。右边是正在形成的“子代”长绳。复制过程从右往左进行。结束后，两条子代长绳将具有与父代同样的化学结构

两条核苷酸序列的精确互补意味着两条长绳上携带的是同样的信息；如果知道一条，就能重构出另一条。互补性为精确复制DNA分子提供了途径。图9.10展示了一小段DNA双螺旋。在复制过程中两条长绳会解开，露出核苷酸碱基。DNA聚合酶会生成新的长绳与已有的长绳互补。当过程结束，就会形成两条DNA双螺旋，每一条都与父代一样。

不通过这个模板机制，细胞就无法生成新的DNA分子。生成无模板的DNA序列对于细胞将是灾难。随机生成的DNA序列只会产生垃圾信息，理由很简单，通过随机方式生成的东西有用的概率微乎其微。

细胞中还有其他结构也只能通过生长或复制已存在的结构形成，但膜和DNA已足以阐明这一点。在中学生物课上我们学过细胞通过分裂繁殖。大细胞分成两个较小的细胞，然后继续生长，直到可以再次分裂。细胞分裂的基本特点是重要结构的复制体或片段会分到两个子代细胞中。如果不是这样，细胞就无法再生。细胞作为一个整体不会像噬菌体那样通过重新组装产生。因此生物学家只能想象出一个无穷无尽的细胞分裂链条。所有证据表明这个链条会从现在的每一个活细胞回溯到一个共同的始祖细胞，也许更准确地说是唯一的始祖细胞群。

地理和化石证据表明这个始祖细胞群生活在至少35亿年前。原则上现在所有的生命不必都是源自同一群细胞，但所有已知细胞都具有许多共同的特性，这意味着存在一群拥有这些共同特性的始祖细胞群体。除此以外任何解释都很难自圆其说。仅列举少数所有现代细胞都共有的特性：基因由DNA组成，相同的遗传码，信使RNA，核糖体，基于磷脂的膜，大量生物化学网络特征，以及许多蛋白质具有类似的结构和功能。这个共同的始祖已经非常复杂，因此其本身也不可能是地球上的第一个生命。

细心的读者可能注意到除了信使RNA和tRNA之外，我们没有提到RNA在细胞中的各种其他角色，也没有提到另一种作为生命标志的大分子多糖。RNA和多糖对于细胞都很重要，但基于不同的原因它们没有被包括在我们的简要说明中。多糖是生物化学网络不可或缺的一部分，而RNA分子则与蛋白质互动或发挥类似蛋白质的功能，除了信使RNA，它们的功能大致可以被包括在蛋白质相互作用网络中。一些人可能会不

同意这个论断，但将纯RNA子网加入前面讨论的超级网络中并不会明显改变我们总体上将细胞视为复杂网络的观点。

那么细胞是什么呢？根据这里给出的观点，它是由三部分组成的网络，由生物化学、蛋白质和基因调控这三个子网深深地交织在一起组成。这个超级网络位于一个有结构的和化学的环境中，可以维持和繁殖。这个超级网络具有惊人的特性，它能够自我指令和自我维持。指令编码在DNA的核苷酸序列中，涉及这个存续结构的方方面面，但是没有说明它是如何开始的。地球上存在大量这样的网络（超过10^{20}）；并且每个生命（即网络）原则上都能回溯到至少存在于30亿年前的一群网络（即细胞）。

超级网络的节点和边包括DNA序列、蛋白质、蛋白质与其他成分的相互作用、化合物以及化学反应。DNA序列的合成和表达受蛋白质控制，还有各种化学反应也受蛋白质控制。甚至作为生物化学网络节点的各种化学物质出现在细胞中也是因为蛋白质酶和跨膜传输蛋白质。蛋白质由DNA序列决定，DNA序列反过来又被蛋白质控制、修复和复制。这个网络极具可塑性，我们身体里细胞类型的多样性，以及自然界中生物和生物行为的多样性，都证明了这一点。所有这些都源自复杂网络与环境的互动。

刻画生命的网络的一个基本特性是，许多甚至绝大多数节点和边单独来说对于网络的存在都不重要。当一个新的基因在细胞中表达，就会增加了新的节点和边；当基因关闭，一些节点和边就会消失，但网络作为一个整体会保持稳定；只有细胞死亡它才会崩溃。据我们所知，从所有现代细胞的最初的共同始祖开始，这三个子网络就已经契合在一起。在具有多种细胞类型的多细胞生物的细胞中，不同细胞的网络特性会有显著区别，生物的DNA中的基因只有1/3到一半会在一个特定的细胞中表达。从某种意义上说，身体中的细胞通过合作组成了一个网络，每个细胞都是这个更大的网络的一个子网。每个细胞网络和身体层面的网络可以回溯到单个细胞，受精卵，以及之前未受精的卵子和精子，再往前是其他身体里的细胞，不断往前追溯直到最早的始祖细胞。这个连续的历史因果链条是地球上的生命的一个基本面貌。

复杂适应系统是什么？

生物体为复杂适应系统（CAS）提供了最基本的范例。其他例子包括国家经济、蚁群、生态和人脑。所有这些系统都在面对变化的外部事件时表现出一定程度的稳定性。在CAS研究中重要的概念包括涌现、自组织、适应、自稳机制、通信和合作[3]。所有CAS都可以被视为网络进行研究，但不是所有的网络都具有适应性。CAS与其他复杂系统的不同之处在于节点具有记忆。有了记忆就可以在响应时考虑过去发生的事件。在对这些系统的研究中，节点通常被称为“智能体”（agent）。智能体可以很简单，但不能过于简单。沙堆会表现出复杂行为，但不是CAS；太阳系和雷暴也都不是CAS。沙粒、无人居住的行星以及大气分子都过于简单，无法产生CAS的适应性行为特征。

复杂适应系统的特点是节点（智能体）基于过去的记忆和新的输入进行决策。所有已知的CAS的节点和边都是预先设定的，并且如果不是全部至少绝大部分都可以描述为具有调控边的网络。不经过仔细的预先设定，随意组合的节点和边自发表现出CAS行为的可能性微乎其微。生命网络中的蛋白质及其互动，大脑中的神经元，生物圈中的生物，以及股票市场中的交易者是具有记忆的预先设定节点的一小部分例子。

设定的说明需要指令，由于复杂引擎是目前所知的唯一能生成复杂指令的机制，因此最终是复杂引擎使得CAS成为可能。有时候，复杂引擎也在其中运作。下面我们将探讨两个典型的CAS系统——人类大脑和社会组织。

第10章　人类学习和创造性

如何理解人类思维

想象自己在热带海滩，蓝天白云，风吹拂着棕榈树，海浪拍打，白色的沙粒按摩你的脚趾。这很容易，只要到过海滩就能想象。同样的能力使得数学家能证明定理，雕刻家能从一块大理石构思完成的雕塑。理解思维是最大的智力挑战。自从有了文字，人类就开始思考这个主题，一些最伟大的思想家对这个问题也望而兴叹。通过研究我们头部的这个器官——大脑，科学家们才开始在对思维的理解上取得进展[1]。

生物学家们一致认为大脑同心脏和胃一样，是一种生理器官。虽然我们对它还没有完全研究清楚，它的功能却是没有疑问的。人体中所有的组织器官都有特定的功能，所有这些功能都是基于大量细胞的协作。大脑的主要功能是信息处理，这一点几乎没有疑问。过去半个世纪以来，对于这个功能是如何执行的已经了解了很多，但还有一个问题：这个信息处理器官是如何产生出自我的意识感即“心智”的？一些人认为这种硬件中心观是一种冒犯或威胁，但不这样就只能停留在蒙昧的神秘主义。固守意识的非生物解释只会导致将人性从动物根源剥离的古老愿望——换句话说，认为人类具有特殊地位。坚定地采取科学立场并不意味着藐视这个问题。大脑到底是如何产生出自我意识的，是科学最大的未解决的问题之一，但神经学家们基本都认为，这个问题终将被物理定律以及计算和数学原理解释，并且不太可能需要新的物理学。

意识让我们能将当前的感官输入与记忆结合起来，从而生成连续的情境。这反过来又使得我们可以根据我们感知到的自身所处的情境进行决策。意识让我们能领会复杂和抽象的问题，并在内心比较各项行动或解释的优劣。意识也让我们能假想采取各种行动。引用卡尔·波普尔的话说，意识“最终使得我们可以通过假设以代替我们的死亡”[2]。

但意识并不是这一章关注的重点；我们的目标是理解学习和创造性。大脑研究的一个基本问题是我们是怎样学习新事物。我们的大脑能够编码过去的事件以及关于这些事件的结论，我们为何能对之前所学的

进行扩展？我们已经看到复杂引擎具有内在的创造性，很善于解决基本不可能解决的难题，即解空间非常庞大，而且其中大部分解都没有用的问题。大脑一直都在解决这样的问题。那它有没有利用复杂引擎呢？

信息在哪里？

大脑由神经元和其他细胞组成，神经元之间通过突触传递动作电位，也通过分泌化学物质向周围释放信号。从信息的角度看，扩散的化学信号携带不了多少思想内容的信息，因为化学模式很简单，通常是以梯度的形式提供位置信息，或是改变附近神经元的行为。可能的化学模式的数量很有限，远远比不上可能的突触模式的数量，因此能够编码的信息量也少得多。所以重点是动作电位。

我们可以认为神经元传递一次动作电位可以传送一个比特的信息。这样每秒能传递数千次动作电位的神经束就能每秒传递数千比特的信息。这似乎很合理，但证据却明确否定了大脑以这种方式编码有意义的信息的观点。

分析神经元控制肌肉收缩的过程，有助于我们了解动作电位如何传送信息。当大脑指挥身体运动时，会调整不断冲击我们身体里每个肌肉细胞的动作电位。对于一些肌肉细胞，动作电位的频率会加快，导致这些肌纤维收缩，另一些的频率则会减慢，导致其舒张。当身体运动时，大脑不断调整流向肌纤维的大量动作电位，使得数百万肌肉细胞协同收缩和舒张。大脑就好像在操纵有数千根线的木偶。动作电位流就好像这些线的拉力。指挥肌肉运动的信息不是编码为单个动作电位，而是动作电位流频率的加快和减慢 [3] 。大脑中的模式可能也与此类似，只是扮演关键角色的是更复杂的模式。

动作电位活跃在整个大脑，没有哪个神经元会静默很久。部分动作电位是神经元自发产生，但大部分是受进入大脑的感官刺激引发，以及来自大脑其他部位的动作电位引发。所有冲击神经元的动作电位都通过激励或抑制型突触传递。接收端的神经元就好像微处理器，每过几毫秒就对输入进行整合，一旦一定时间间隔内激励型输入高于抑制型输入就产生新的动作电位。神经元之间的连接模式使得动作电位不断在3维的时间敏感网络中循环。抑制型突触和生理结构对电位流进行约束。正反

馈环则让特定的模式可以扩大响应区域。在任何时刻活跃的群体都包含数百万神经元。短暂存在的活跃群体不断有新的神经元加入，然后又被新联合起来的神经元群体替代。

局部网络可以与其他网络互动。节奏很关键，因为动作电位到达的时间间隔超过几毫秒就不再有累积效应。突触的连接强度可以改变，神经元的阈值也可以改变；通过加强或削弱神经元之间的连接，可以调整它们的相互影响。同时，新的突触可能产生，老的突触也可能消失。人们普遍认为学习是通过加强或削弱突触连接巩固特定的模式。增强的模式形成记忆。增强的突触使得特定的模式在未来更容易再次出现。

组成思维或思想或内部模型的动作电位流的3维变化模式的细节必然极为复杂，高度并行，具有冗余，并且充斥着噪声。思维随时间变化，并且每次出现时的细节都可能不同。思维不可能还原成特定时刻在网络中传播的动作电位比特的总和。信息必然以某种方式被编码为动作电位流的强度、节律等模式。这一点可以用更简单的系统来类比。例如，消息可以用莫尔斯电码编码并通过电线传播，关键不在于通过电线的电子数量，而在于电流的启停模式。类似的，说话时传递的信息内容不是由空气振动的次数决定，而是取决于振动的模式。

理解思维的信息载体的一个主要障碍是信息在大脑中的编码模式截然不同于计算机的01序列和人类说话时的线性词语流。很显然，不同的连接模式可以表示不同的意义。但要弄清楚3维的细胞交流模式如何发挥作用，使得神经元的输出可以正确调整邻近的网络从而形成有用的新网络，并不是那么容易。尤其是要让新的网络能产生正确的肌肉运动序列，或正确的视觉，或能正确传达意义的声音序列。另外也缺乏概念工具帮助我们正确理解神经元组成的3维网络如何编码特定的信息。用之前讨论的术语来说就是，神经元交互模式是如何实现指令的？目前神经学家还不能完美地解答这个问题。

总的来说，无论是指挥肌肉运动的信息，还是表示思想和思维图景的信息，或是表示大量学习的知识的信息，都嵌入在动作电位流组成的3维时间敏感网络中。这种概念层面的描述很模糊，但除非某些重要的事情被神经学家忽略了，否则大脑的运作原理或多或少就是这样。这个模型没有告诉我们某个具体的功能模式是什么样子。不过这个模型要能有任何作用，我们还必须在大体上解释神经元连接功能模式是如何形成

的。

如何进行决策？

考虑一个问题，假设你知道一个6位字符串的前5位，比如10010，你如何预测第6位？有3种可能。第一，可能存在某种规则决定下一位。如果你知道这个规则，你就能有100%的把握预测下一位。第二，下一位可能是随机的，也许是抛硬币的结果。如果是这样，你的预测最多只能有50%的准确度。第三，下一位可能是规则和随机的某种组合。如果是这样，如果你知道规则，你预测的准确度就能大于50%，但低于100%。计算理论没有给出其他可能。

由某个人“想出”下一位不是一种计算选项，不过仔细审视一下“想出”0或1是什么意思可能会很有趣。也许在我们的头脑里有一组规则，在相同条件下会产生一样的0或1；或许我们能通过思维“抛硬币”来作出决定。显然，大脑也能够以某种方式结合两个极端，在重复的情形中产生不可预测但是有偏的结果。

研究这种问题的心理学家面临的困难是，不可能通过实验检验在我们的头脑里到底产生了什么想法。这是因为在对人类进行测试时不可能重复完全相同的条件。对人来说，测试的条件除了给出的任务之外，还包括受试者在想什么，而受试者所想的显然会包括他们之前参与的测试。在第一次与第二次测试之间发生的一切会不可逆转地改变受试者的心智状态。因此，不可能重复完全相同的条件。我们知道人类并不总是可以预测，甚至有时候不讲逻辑，因此我们的思维功能不可能是完全确定性的。另一方面，我们又很不擅于生成随机01序列，因此应该也可以排除我们的思维是无偏的随机数发生器的可能性。这意味着我们可能采取的是某种组合策略，我们的思维过程不是随机的，但也不是完全确定性的。这可能具有进化意义。很可能在进化历史中，行为上随机性和可预测性的混合能够在平均上产生最成功的结果。记忆使我们能利用过去的经验，随机性则在我们的行动中引入不可预测的因素。也许这就是我们说的创造性？在面对对手时，不可预测性能带来很大的优势。在被狮子追击时，如果黑尾羚羊总是在狮子距离10米时左转，距离5米时右转，狮子不久就能预测出它的转弯，然后只需直奔终点就能抓住猎物。在生命的许多情形中不可预测性（或者说，创造性）都能带来优势。

在做决定时，我们利用记忆存储的信息以及感官接收的信息。随机波动提供了输入的第三个来源。如果我们躺在一个完全黑暗无声无味的房间里，让我们的思绪漫游，除非我们直接就睡着了，否则我们很快会产生新的思维；我们的思想也不会限于记忆的事件。新奇思想的信息从何而来？内省意味着新的思想来自记忆事物的重新排列和组合，但这并不意味着它们只是旧记忆的简单重演。创造性的本质意味着这种记忆信息的“翻弄”不是完全预设的，其中必然涉及随机试验的因素。毕竟，创造性的定义就是产生不可预测性。

记忆碎片的随机组合不可避免地会产生许多垃圾思维。有些思想对面前的任务没有用，甚至很可能哪都用不上。要理解我们如何思维，很重要的一点是要搞清楚我们如何聚焦于有用的思维。

我们的世界，以及我们的思维，是确定性的吗？

从伽利略开始，400年来科学家们一直在稳步推进我们对世界的认识，并建立了这样的观念，即我们的宇宙是4个方面在时间中互动的产物：离散性（即基于物体的宇宙固有的可列举性）、几何、规则和随机性。我们所处的环境大部分是离散的：存在一个个的物体。由于宇宙中到处是可区分的事物，计数和数字成为可能。如果一切都是平滑的流动，怎么可能会有计数呢？数字将毫无意义。

规则有两种形式：逻辑和物理定律。逻辑是数学的基础，幸运的是，物理定律符合逻辑。据我所知，还没有人从科学的角度解释过为什么宇宙是逻辑的，或者不符合逻辑的宇宙是什么样子，或者不符合逻辑的宇宙是否有可能。当逻辑应用于在几何空间和时间中具有特定性质的对象系统时，就必然存在规则性，并且规则性可以用数学描述。这就是我们所谓的物理定律。

宇宙中也到处都有随机性。混沌理论从数学上证明了，在一些情形下，即使系统的数学描述是完全确定性的，也不可能精确预测长期的结果。这种系统的特点是对关键参数极为敏感，如果要计算未来任意时间的结果，就必须有无穷精度[4]。一个例子是天文学中的“三体问题”，相互之间有引力的3个天体相互围绕运动，组成了一个混沌系统。时间越长，对3个天体未来位置的预测就越不精确，虽然相关的物理定律和关

系的数学描述都很清楚并且很精确。难点在于无法以足够的精度确定3个天体的质量和速度，因此也就无法在相互围绕许多次后还能准确预测它们在空间中的位置。预测的时间越长，不确定度就越大。这种不确定度就是所谓的确定性混沌。包括气体和液体分子的运动在内，许多系统都有这种特点。

混沌是看起来随机但数学描述可以很精确，量子力学则完全建立在随机的基础上，微观事件具有内在的不可知性，只能用概率表示。有两个例子展示了这种随机性，一是特定的放射性原子的衰变时间有本质的不可预测性，二是不可能同时确定一个很小物体确切的位置和动量。

现代量子力学的创建者之一阿瑟·康普顿给出了一个非常清晰的例子：

现在想象一束微弱的光线通过一个小洞，然后由于衍射扩展成很宽的一束光。在这个宽光束的光路上我们放置两个光传感器，A和B，分别连到放大器。光感器非常敏感，只要进入一个光子就能记录到。光路上有一个快门，打开的时间刚好足以通过一个光子。这个光子会进入哪个光感器？我们无法确定……第一个光子可能进入其中一个光感器，在初始条件完全一样的情形下，下一个光子可能进入另一个光感器。因此结果是不可复制的。就我们所知，这完全是随机的[5]。

由于单个原子或分子在遇到光子时运动会改变，因此每个原子都是某种光感器。一旦光子冲击气体或液体，导致的分子运动的变化就有内在的不可预测性，因为无法知道一个特定的光子会影响哪个原子。量子力学的不确定性与确定性混沌结合在一起，使得我们的宇宙充满了细节不可预测的随机原子运动。这并不意味着发生的一切都是随机的——远不是这样——而是说随机性永远无法完全排除，并且总是存在于分子层面。

当我们想象一个由宇宙中概念上可能的所有事物组成的假想空间时，随机分子运动的存在意味着只要有足够的时间，任何局部的分子形态都有可能在某个时刻出现在某处。而如果没有随机性，许多情形将永远不可能。如果我们的宇宙是确定性的，则未来的一切在最开始就已经确定了，所有的行为，甚至我们的思想，都将只不过是一个巨大的自动机的推演。这种情形中的概率反映的只不过是无知而不是不确定性。相

反，当系统中引入随机性，概率反映的就不仅仅是无知，还包括了不可知。好的一面是具有随机性的宇宙也允许创造性，没有随机性，真正的创造性就是不可能的[6]。

回到大脑，实验证据表明大脑内部的交流是“噪声”。一些神经元同时激发，当实验探测同时激发的神经时，个体细胞的表现并不一样。如果个体细胞的特定活动源自分子层面，分子层面的不确定性就必然影响细胞的行为。有了这样的不确定性，如果大脑是确定性的，将非常让人惊讶。更有可能的是大脑具有概率性。

大脑功能的一般原理是什么？

成年人的大脑据估计有1000亿个神经元，每个神经元平均有大约1万个与其他神经元的连接。因此大脑的神经元连接的量级是1000万亿（10^{15}）。突触在神经元之间传递动作电位。在传递的链条中，每个神经元就好像一个简单的微处理器，在很短的时间间隔内，如果激励型突触输入压倒了抑制型输入，就会产生一个新的动作电位。一些神经元还能自发产生动作电位。至于什么时候产生动作电位的“决定”则是一个简单的（化学）计算。神经元响应刺激时的动作电位的传递通常沿着复杂的多条路径，取决于连接的几何特征以及参与的神经元的协同行为。在互连的网络中大量神经元的相互作用的结果极为复杂，绝对算得上是计算[7]。在实验室中可以培育单个神经元，这样可以研究两三个或者少量神经元之间的简单通信模式；而在人脑中，回路经常涉及数百万神经元。

大脑结构是模块化的，特定的神经元群执行专门的信息处理功能。局部的专门计算组合起来完成更大的任务。一个全局任务是处理和整合每时每刻感官细胞产生的大量原始数据。我们的感官——视觉、听觉、触觉、味觉和嗅觉——以及广泛的内部监控系统，都依赖数百万感官细胞不断向神经元传送动作电位来报告它们的状态，然后这些动作电位又被传送到大脑的特定部位。神经系统的一个主要功能是过滤巨量的原始数据流，识别出未来可能有用的信息，将其他信息抛弃。这项繁重的工作，主要是无意识进行的。大脑执行的另一个主要的计算工作是协调对输入数据的反应。这主要是协调肌肉的活动，让我们可以登上公共汽车或用手而不是额头抓住棒球。肌肉是通过动作电位进行控制，动作电位

从大脑沿神经传递到肌肉细胞，引起它们收缩或舒张。大脑就像一个控制中心，介于感官细胞和肌肉细胞之间，监控着大量源源不断输入和输出的动作电位。大脑同时也是记忆库。学习可以是有意识的，也可以是完全无意识的。

在对感官输入进行响应时，将其与过去类似的输入数据模式以及相应的后果关联起来会很有助益。甚至大脑极为简单的动物都会有基本的记忆。动物也有欲望。欲望可能是简单的“寻找食物”或“交配”，也可以复杂而多面，就像人类的互动。记忆和欲望会影响大脑产生的结果。另外，人类还能构造内部的世界模型。我们生成思维图像，看到物体背后的东西，“听到”头脑里的声音，“感受到”想象的行动的后果。这个内部的模型世界让我们能试探可能的行动步骤，而不用实际运动肌肉。一旦想象的后果被认为可行，就会触发真实的行动。

记忆肯定是一个选择过程，我们不可能记住感官细胞的所有输入，这样做也没有意义。大脑的一大任务是从感官和内部模型世界中识别出“重要”的模式。当足够重要的模式被识别出来后就会被提交给记忆。记忆意味着信息的存储。大脑的特殊设计使得有可能形成大量的状态，即神经元之间的通信模式。

任何能有多种状态的物理系统都能编码信息，这也同样适用于大脑功能。大脑是一个物理网络，用活跃的子网表示外部世界的各方面，因此特定的信息必然是编码为子网的各种状态。

离散元素组成的有限系统（即由部分组成的系统），可选的状态总是有限的。经常是一个系统的状态与另一个系统的状态互动。通过这种方式，一个系统能对另一个系统产生多样而又特定的影响——比如司机（一个系统）与汽车（另一个系统）的互动。互动能力的大小也就是信息量。无论是人类还是动物的大脑，都存在神经细胞（神经元）之间的互动。基本的通信机制是第4章讨论的动作电位。大脑各部分局部的通信模式之间产生互动，这些互动又引发其他模式的建立。这个总体上的认识很显而易见。难点在于具有特定意义的子网是以何种方式建立起来的。

由多个物理对象组成的3维系统，各部分之间的互动遵循物理定律，其中的信息内容的分析可能非常复杂。因此大部分信息学家只关注

线性符号序列携带的信息。01字串很符合这个要求，幸运的是，我们可以证明更复杂的系统所编码的信息总是可以转化成适当的01字串而不会损失意义。因此适用于01字串操作的计算原理也适用于更复杂的系统，其中就包括神经元的通信网络。

什么是进化认识论？

认识论是研究知识的起源和局限的哲学分支，关注的问题是我们是如何知道我们所知道的。1974年，心理学家唐纳德·坎贝尔在介绍卡尔·波普尔关于人类学习的研究的文章中提出了进化认识论一词 [8] 。计算的视角带来了新的认识。它彰显了一个重要的事实，即计算机能复制另一台计算机存储的内容，而我们人类却无法复制另一个人的记忆。我们能向他人学习，但我们无法精确复制他人头脑里的东西。这提出了一个严肃的问题，就是在我们学习的时候到底发生了什么。

如果你想让你的计算机能够“知道”其他计算机已经“知道”的东西，你只需要将相应的文件从一台复制到另一台。复制的是编码了所需信息的准确位串。如果一个人知道某件事情，相关的信息被编码为大脑里的神经活动模式。如果两个人知道同样的事情，他们的神经模式并不是一样的。当我们向老师学习时，我们不是复制的她大脑里的神经连接模式，我们是在创造我们自己的能表达相近意义的模式。类似的，当我们读书时，我们不是将记忆的信息编码为大脑里的字符序列；我们是在每个人独特的大脑中创造能表达同样内容的神经连接模式。

因此“知道”的核心问题是我们的大脑是如何创造出特有的神经连接模式，使得各人的大脑中不同的连接模式传达出同样的意义？要回答这个问题我们需要更深入地了解什么是心理表征。本质上，当我们产生了一个想法或学会了某种新事物时，在我们的头脑里有了怎样的不同？

如果你将自己缩小到细胞一样大，一路追踪动作电位在神经元之间的传递，你会发现有些路径会回到起点，这表明存在反馈回路。有些路径则会前往大脑的其他区域，与其他反馈环路和数据流互动。大脑中的神经元不是随机连接的，但也不是预先就设计好连接的所有细节。最初的模式显然是由总体规则决定，许多细节则待定；第7章曾对此进行探讨。在我们出生后学习的过程中，有些连接被添加，有些则被删除。结

果是所有人类大脑都在整体上相似但细节却不同。这意味着两个不同的大脑在接收完全一样的输入时，导致的动作电位的传输路径和模式不会是一样的。这与计算机很不一样，如果输入相同，两台相同型号的计算机产生的电压和电流模式也是相同的。

从大体上说，大脑的核心活动就是“认识”个体的当前状态和周围环境，并基于这种认识行动。“认识”在这里指的是在当前的情形、之前情形的记忆、之前经验的概括、未来可能的行为和后果之间建立有用的关联。概括是学习的核心。经常会有一系列事件表现出共性。与将每次经验作为没有关联的条目进行记忆相比，总结出共性再记忆要高效和实用得多。人类大脑特别擅长寻找关联各种情形的模式。对这些模式的识别和记忆也就是对概念的学习：概念使得意义可以在不同的人之间交流，让不同的人可以对同样的现象有相似的认识。

无论是感官输入还是概念输入，通常并不立即形成“认识”。它并不会马上被记录为有用的形式；而是要过一段时间。这之间发生了什么？没有人确切地知道，但有些事情是可以肯定。首先，当我们意识到某事时，这种意识必然是表现为神经元之间的动作电位交流模式。由于思想进入意识的时间不止一瞬间，因此这种活动模式必然具有某种程度的稳定性，而这只有信号以某种反馈的模式不断循环才能做到。著名神经科学家和科普作家，诺贝尔奖获得者杰拉德·埃德尔曼称之为“折返”[9]，在具有往返连接的大规模并行神经结构之间持续的动态信号交换。

我们可以肯定的另外两件事情是，连接模式是动态的，并且连接模式与一部分大脑对另一部分大脑的意义有关。我的意思是它们必须相互配合，动作必须相互协调，并且在更宽泛的背景中能产生有用的后果。这其中有许多问题。动作电位在局部传播的模式必须与其他局部的连接模式交互，产生出一致的全局模式，组合成某种有意义的东西。局部网络之间的交互必然会激发神经元“交谈”的新模式，从而又与其他正在进行的模式交互，不断推进。连接模式要相互配合（即在更大的背景下有意义）才能产生全局行为（包括思维），要能在生物体的当前状态、外部环境、欲望以及相关记忆的背景下产生有用的结果。这些看似不可能的复杂要求集合到一起，将我们带回了这一章的中心问题：如何才能产生能与记忆和期望在内的各种正在进行的大脑活动有效配合的连接模式？我们知道，在我们每次遇到新的事情，或想要某个东西，或学习新的知识，或产生新的思想时，大脑中并没有一个天才工程师在设计新的

连接模式。

现在有哪些理论？

在第6章我们探讨了人工神经网络（ANN）。ANN可以用计算机编程实现，并表现出惊人的学习能力。ANN受神经生物学启发，研究过大脑的人都会同意大脑功能是基于神经元组成的网络。ANN的学习是通过调整权重，权重决定了输入对节点的相对影响力（见图6.1）。在大脑中，权重由突触的强度和性质决定。在计算机模型中有许多方法可用于权重的调整，这些方法都需要训练。其中一个方法是对权重进行细微的随机变化，如果变化能改善性能，就保留作为下一轮调整和选择的基础。采用这种策略时，学习（能改进性能的权重调整）是源自累积的选择——进化过程。训练样本越大，结果越好。另一个方法是反向传播，是确定性的。这个方法需要有期望的ANN输出，并且每个节点的输出能表示成数值形式。如果满足条件，就有可能写出方程计算各权重值的变化，使得对于特定的数据集，给出的输出与期望输出之间的差别最小化。训练是针对不同的数据集反复应用反向传播。得到的权重是各训练集的优化值的折中，代表了对训练的数据集的归纳。两种方法学习的权重集通常都能很好地处理ANN之前从未遇到过的输入。

一些学者分析了大脑采用反向传播产生正确的连接模式的可能性。所有研究都表明可能性极低。神经元的确有可调的阈值和可调的突触强度，突触强度的增加和削弱很有可能是学习的重要方面。但很难想象神经回路的输出能表示成数值形式；同样也很难想象怎样的生理结构能执行所需的数值计算并调整数百万突触的强度以实现结果。可能还需要另一个脑来专门负责调整参数。大脑中并没有这样的并行结构。

近年来一个名为“贝叶斯脑”的新理论越来越受关注[10]。在这个理论中，大脑也被视为神经元网络，但网络编码的是概率分布。概率分布给出所有可能结果的可能性。在这个大脑概念中，感知不是所看见的。当我们“看见”一棵树时，我们不是在从眼睛直接提取感官输入，而是在感受我们的内部世界模型。眼睛输入的数据不是被有意识感受，而是被大脑用来调整我们的内部世界模型，因此我们的思维之眼所“看见”的密切反映了在那里被风吹动的实际事物。我们所看见的与实际在那里的事物之间的匹配并不完美，但大脑在不断调整其内部状态，将内部产生的

数据（思维之眼）与感官产生的感知数据之间的不匹配程度最小化。统计学家称之为最小化不确定性。这种理论源自对人类视觉的研究，视觉信息的处理目前仍是实验主导的领域。

这个理论有几个部分。首先，更高级别的大脑核心发展和维护关于我们所处世界的不同方面的思维模型，并用这些模型不断评估和预测。第二，预测和感知数据都编码为概率分布。这意味着数据和预测都包含对其不确定度的估计。第三，感知数据处理的主要任务是最小化不确定度，让输入的感知数据与预测相匹配。模型预测与输入感知数据之间的不匹配（不确定度）可以从两个方面减小，要么是改变感知数据（例如移动），要么是修正内部模型。修正内部模型是学习的一种形式。

这个理论之所以最近很受关注有两个原因。首先，在用实验测试人们对心理测试作出的反应时，他们的反应接近“贝叶斯最优”。这意味着我们是基于一种统计推断进行决策，这种统计推断是托马斯·贝叶斯在1761年提出的，后来逐渐发展成统计学的一大领域。其次，当贝叶斯统计应用于连接编码了统计分布数据的网络时，不确定度的数学形式与物理中的自由能相同。这意味着大脑模式的变化可以用物理学家分析物理系统的相同数学方法进行分析。自由能是物理系统在改变时最小化的热力学参数。类似的，贝叶斯大脑理论预测大脑的变化受不确定度的最小化驱使。卡尔·弗里斯通称之为将“意外”最小化[11]。

在神经网络中，如果节点有关于其权重的不确定度的信息，最小化不确定度将最大化输出的精度。如果输出是预测，不确定度反映的是预测和感知输入的不匹配，则通过调整网络权重减少不确定度就可以改进模型。因此通过将“意外”最小化，大脑就能自动调整其内部模型，使得内部预测与感知数据匹配得更紧密。

这个理论很引人注目，因为它似乎解决了大脑优化编码世界模型的网络权重（突触强度）的问题。它也有让人疑惑的确定性问题！在其最简形式中，贝叶斯脑的概念似乎意味着编码世界模型的网络“自动”调整以优化与感知数据的匹配。贝叶斯脑是一个数学模型，必须用某种物理设备实现。目前的尝试包括某种形式的层次性大脑组织，高层大脑将内部模型的预测向下投射，低层大脑则计算预测与输入感知数据的不匹配（不确定度）并向上投射。根据对不确定度的评估对连接强度进行修正，使得改善后的模型产生出的预测能与输入的感知数据匹配得更好。

这是一个迭代过程，反复循环打磨内部模型，以产生出与数据能充分匹配的预测。计算机模拟实验取得了较好的结果，但是当模型与感知数据不能紧密匹配时会很困难。至少会出现两个问题。首先，自由能原理并没有说明该如何设计一个能建立全新的内部模型的系统；其次，复杂网络几乎总是会陷入复杂性科学所谓的“局部最小”。局部最小在大脑中就是对当前最优值的任何改变都会导致不确定度的增加（数据不匹配度的增加），然而在参数空间的更远处却存在表现更好的配置。贝叶斯大脑是数学模型而不是物理模型，虽然从理论上大脑有可能是这样运作的，但还需要有具体的方法来实现理论，并与所知的大脑知识相容。目前计算机模型已得到了一些进展，但最后的结果还是未知数。

还有一个似乎很不相同的策略也很适合解决开放式问题，并且能避免局部最小问题。这就是复杂引擎。它也是需要物理证明的抽象思想。复杂引擎在大脑中的实现需要有多个假说（模型）以及随机混合其他模型要素的机制，这些模型与当前的问题并不一定很相关，可能还需要有添加或删除“随机”网络特性的机制。大部分备选假说的预测都很糟糕，但有一些可能有希望，进一步提炼有可能最终得到比目前最好的模型更好的结果。这样的策略需要在解决问题的过程中，大脑同时进行多个并行尝试。我们在第5章对复杂引擎的分析中看到，要高效地实现这个策略，每一轮都需要产生多个改进的方案，而且随机或半随机的改变的尺度要适中。研究大脑的人都知道大脑皮质的大规模并行性。这种结构似乎很适合执行复杂引擎的计算所需的并行处理。

复杂引擎在大脑中可能起作用，有两个原因使得这个想法很具有吸引力：（1）以往的经验表明，只要应用得当，复杂引擎型的计算很擅长为困难的问题找到很聪明的以前从未想到过的答案；（2）这种方法与贝叶斯脑的原理相容。一旦复杂引擎的思想与贝叶斯脑相结合，有一种可能很快浮现出来。最小化意外能为备选假说的选择提供标准。还有一条不那么明显的思路是，如果模型预测与输入数据之间的不匹配（意外）很小，只需利用大脑自由能原理就能简单实现参数的快速调整（通过反馈改变突触强度）。而如果这个快速过程不顺利，输入与模型预测之间的不一致程度一直很高，可能就会触发复杂引擎策略进行范围更广的尝试。后面这个过程很有可能就是当人们在“思考”某件事的时候发生的事情。

这是大脑研究让人兴奋的时刻。贝叶斯脑与自由能原理为大脑运作

原理的研究建立了一致的框架，并不断带来让人兴奋的新发现。

我们如何发明？

对于大脑在神经元回路层面是如何学习还有许多不确定的地方，而在人类体验的意识层面上对此则有大量相应的观察。埃德·索贝博士是美国发明家名人堂的首任董事和美国西北发明中心的创立者，他在全世界宣扬人人都可以更具创造性的思想。他有一项本领是可以让一群人当场发明出某种出人意料的东西。他认为，所有伟大的发明家都是通过试验和试错发明[12]。爱迪生就是著名的例子。他的名言“天才是1%的灵感加99%的汗水”说的就是这个意思。当面临一个答案不明显的问题时，发明者会进行尝试，检验较好的想法，进行修改，再检验，再修改，不断反复。在一个真正好的设计方案出台之前，往往经过了数百次的原型设计。许多发明家都说自己是“用手思考，”意思是他们在解决问题的时候，会动手将材料拼凑出原型，操作，修改，测试，再修改。“动手”的过程给了他们用更好的材料再次尝试并修改原型设计的灵感。有时候最终的产品与最初的尝试一点也不像。索贝博士很清楚，最初快速而粗略的尝试加上后来不断改进的环节，这样的发明方法比通过长时间思考想一蹴而就的方法要有效得多。在对这个发明过程的描述中没有明说的是，在每次实际尝试之前，发明者在头脑里已经考虑并否定了大量可能的修改。

现代科学的原理与此类似。科学是一个累积性的知识体，通过不断反复地提出假说、检验和修改假说的过程建立起来。具有创造性的人会不断提出新的思想进行检验。但凭空想象很少会得出全新的东西。“头脑风暴”是一项成熟的技术，可以让一群人产生出尽可能多的想法，其中许多想法都很怪异，目的是找到一个想法改进后能更好地解决某个难题。社会比个人更擅于解决问题正是因为有许多人提出了各种可能的解决方案以供选择。

进化认识论的一个基本观点是我们的潜意识中运作的也是同样的原理。当面临一个新的情境时，最好的策略——实际也是唯一可行的策略——是基于已有的信息（记忆）进行一些非随机的猜测，然后反复挑选最好的进行修正，直到达到可以应对这种情境的思维状态。在神经元传递动作电位的层面上，这意味着产生出新的神经元通信模式，能与其他

神经元通信模式协调一致，并最终引发成功的行为。

有没有复杂引擎在大脑功能中扮演重要角色的证据？

这本书的核心思想是，一种特殊的计算策略复杂引擎能高效地产生本不可能出现的具有目的性的信息体，而且这种方法在我们的世界中广泛存在。人类思想和人类认知显然包括信息。我们知道思想具有信息，因为它们能够通过说话和书写交流，并且这种交流能用香农和算法信息进行分析和量化。我们对大脑结构的认识也说明在细胞层面上思想就是神经元通信的复杂模式，在大脑中能够产生大量这种模式。

而这些模式中任何一个随机发生的可能性都极低，并且大脑无法预先“知道”如何构造一个符合下一个需求的模式，那么大脑又是如何产生出解决下一个问题的模式的呢？

贝叶斯脑假说从统计的角度来研究这个问题，将大脑视为一台统计推断机器。模型参数的概率分布被视为在模型范围内囊括了“所有可能性”。自由能原理提供了内部生成原则，让系统可以“孕育”出最优解。换句话说，它提供了一种在神经元层面上在各种可能性中进行选择的方法。这个环节很重要，但是还有问题，那就是大脑如何基于之前的模型产生出很不一样的模型。我们如何处理各种我们从未遇到过的问题？

对这个问题有一个很有可能的答案是，新的模型是用之前有用的模型的各方面组合拼凑而成。这个系统（也就是大脑）能通过修改和重组旧的神经模式产生新的模式。从信息科学的角度看，新模式包含的信息必然来自某处，而最明显的来源就是之前的成功模型。不过，这还不是完整的答案。一个成年人大脑中存储的信息显然多于一个婴儿。许多实验都表明我们最初的大脑中都只有硬连线的一些很基本的信息：比如识别人脸的能力，吮吸的本能，以及模仿语言的“天赋”。很难想象一个成年人的复杂思维仅仅是婴儿思维的重新组合。我们的大脑中存储的信息只有3种可能的最终来源：遗传（来自DNA中存储的信息），感官数据的重组，或者对微小的随机变化的累积性选择。

在尝试从计算的角度理解人类学习时，面临的一个基本问题不是表示概念和记忆的通信神经元网络，而是正确的模式到底是如何形成的。因为大脑有数万亿连接，能构造出的可能网络的数量远远多于已知的宇

宙中的原子数量。这使得大脑具有了编码大量信息的能力。但这并不能解决大脑如何创造出能应对特定任务的正确网络的问题。在可能的网络中只有极小一部分能正确地与其他大脑网络相互配合，针对特定的输入产生出正确的输出。自由能原理为选择提供了指南，但要创造出新的结构似乎还存在许多问题。可能的输入的多样性是无穷无尽的。其中包括例如意识到被老虎窥视，对性的渴望，听微积分的课，闻到香味，等等，数不胜数。

最核心的困难就在于我们之前讨论过的概率问题。有许多神经网络都能整合特定的输入并与其他大脑网络正确交互产生出正确的输出，大脑如何找出效果好的那一个？不好用的网络数量似乎占绝大多数。可以想到，感官数据是由硬连线回路采集、分类、增强、分析，实验结果也支持这种观点。但得到的处理信号接下来必须传送给其他脑系统，通过情感、记忆和意识意向对其进行衡量，然后形成行动规划，有时候还要将其付诸行动。这种活动是如此多变，对它们的要求又是如此多样，以至于不可能对它们进行硬连线。新的经验需要新的网络。

大自然在其他场合展示了复杂引擎型的计算很擅于解决大脑面临的这种复杂的开放式问题。如果大脑确实是实现了某种复杂引擎，就意味着其中必然在不断进行无意识的思维和行动片段的变化组合，从而为可能的功能模式的选择提供了素材，以供修改、检验、再修改、再检验，这样不断反复直到能与面临的问题相匹配。在不计其数的可能动作电位网络中，通过反复的匹配和修改，涌现出适应当前需求的那个。在这个过程中，也可能混合进全新的从未出现过的方面，但新的只能在每一次尝试中占一小部分。也许并非巧合，这很像埃德·索贝博士说的发明的过程，但大部分大脑活动都处于无意识层面。要让这样的过程能够工作，就像第5章解释过的，对当前进化的信息结构的每次修改不能太大，否则系统就会因为改进的可能性太小而停滞不前。这解释了为什么我们只有在新的知识与已有的知识和能力紧密关联时才能高效地学习，以及为什么类比在对新事物的理解中扮演了如此重要的角色。

图10.1展示了用大脑领域的词汇重新绘制的复杂引擎。这幅图大致揭示了复杂引擎是如何作用于大脑中累积性的选择网络从而产生有用的结果。这个循环从产生适量的“猜测”开始，即贝叶斯脑框架中针对某个问题的试探性模型（图中的“输出”）。这些猜测以并行的神经通信模式或网络的形式实现。通过将得出的结果与针对的问题进行匹配来检验猜

测。匹配最糟糕的（有最大意外的那些）被抛弃。匹配最好的会被随机修正，产生出多个新的输出。然后又再次检验得到的结果。反复的循环很快会使得得到的结果与问题匹配得越来越好。最后通过“选出最好的”作出决定。如果很紧急，问题又很简单，需要几百毫秒就作出决定。如果问题很复杂或概念很难，则过程可能花费数天甚至数月。许多人在反思自己的成就时注意到，当面临一个困难的问题时，他们需要“睡一觉”，然后答案才会变得清晰。

我对自己在大学学习数学的经历记忆尤深。经常，在面对一个新的概念时，我会首先进行阅读，然后做题练习，但我仍然觉得自己没有真正掌握。过了几个星期，在学习新的知识时，我突然意识到我之前努力学习的概念现在完全清楚了；很容易！这是怎么回事？当时，我对自己头脑里发生了什么没有意识，只是觉得很神奇。

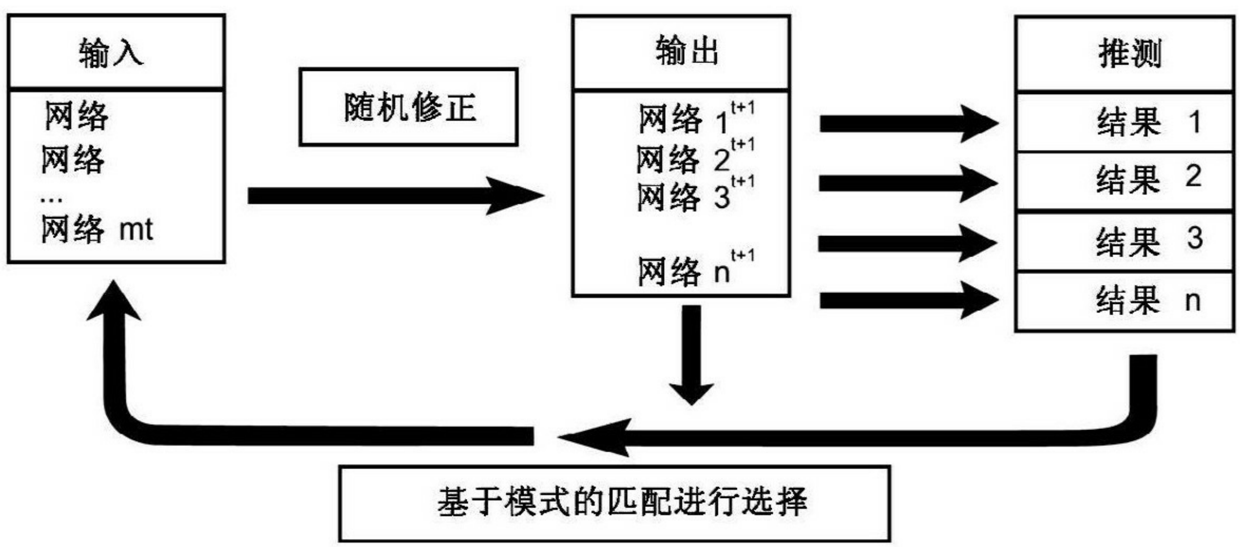


图10.1　根据大脑活动重绘的复杂引擎

当学习是许多人参与的合作行为时，复杂引擎的学习方法变得更明显。这种活动通常属于社会进化的范畴，包括各种以累积文化信息为特征的人类活动，比如科学和技术思想、经济、社会结构、流行音乐和宗教的传播。在第11章我们会对此进行探讨。

总之，复杂引擎在大脑中运作的证据是间接的，说到底，目前还不知道大脑是不是利用了复杂引擎实现其神奇的功能。几乎可以肯定是利用了选择，但还不足以认为就是复杂引擎在起作用。如果是这样就必须

有同时存在的并行解决方案竞争，并且还要有反复的选择、修正和再选择的过程，以利用累积性选择的潜力。复杂引擎是如此强大的策略，大自然在利用强大的解决方案方面又是如此具有创造性，如果这种潜力没有被利用将会让人吃惊，但只有等神经科学家们彻底揭开了大脑的运作原理才能知道答案。

对于人类创造性的来源我们能说些什么？

如果将人类思想和技术视为基于信息的现象，并明确复杂引擎在这种信息的形成中的作用，将能解开许多关于人类创造性的谜团。这甚至似乎能解决古老的自由意志问题。毕竟，在复杂引擎运转时，随机变化为选择提供了素材，结果最终是不可预测的。累积性选择解决了神经网络固有的大量目的性信息伴随的小概率问题。理论分析和实验结果都表明，当选择的素材是由随机变化提供时，引擎最具有创造性。大脑中现存神经网络的随机增删提供了刻画思想的信息的最终来源，并且解释了为什么我们的思想和行为不是完全预先确定的。认为我们的学习是对感觉数据的确定性解释，创造性只不过是幻觉或没有累积性选择的随机变化，这种确定论观点无法解释小概率问题。

复杂引擎解释了在生物中观察到的创造性，确保了任何进化系统的未来都不能明确预测。如果大脑是这样，就为人类的创造性建立了基础。复杂引擎生成的输出固有的创造性同时也解释了我们是如何学习以及为什么我们具有创造性。没有它就很难解释能有效表示各种现象或概念的神经元通信模式是如何克服小概率问题随机产生的。有一件事情我们能够肯定，没有某种形式的随机输入，创造性就是幻觉。

第11章　文化的进化

什么是人类文化？

人类文化与黑猩猩或蚂蚁文化的一个区别在于人类语言。贝多芬交响曲、街角的药店、总统选举、罗马天主教会和棒球比赛都依赖语言。复杂的语言让我们可以与他人分享复杂的思想和知识。从计算的角度看，交流就是发出和接收信息，如果社会成员之间广泛交流大量信息，我们就认为这是文化的一部分。因此文化也就是许多记忆交流的产物。动物也交流，但它们交换的信息量很有限；因而动物的“文化”与人类文化相比起来很贫乏。

人类文化的发展至少有两次重大转变，第一次是语言能力的进化，第二次是书写的发明。这两次转变都极大地提高了文化的复杂度。目前尚不清楚人类是什么时候获得了复杂的语言能力，但进化的过程必定花了上百万年，因为涉及多个大脑区域的大量扩容和喉部的生理构造。这个领域的一些人认为我们现在的语言能力的出现时间只有不到10万年。另一些人则推测复杂的语言是现代人与尼安德塔人区别的标志，并使得我们的祖先具有了决定性优势。人类的智力也与语言能力有关。我们每个人都有许多知识，其中大部分知识的学习是通过语言。我们的许多思维也是以我们使用的语言为框架。

语言使得村庄、传统、手工艺和贸易成为可能。书写的发明使得信息可以在文化中流通，同时也让更多人能够获取信息，从而极大地增加了信息的总量。现在数字革命又再一次提高了这种能力。在书写出现之前，文化信息仅限于能被社会中的个体记住和交流。有了书写，书籍和图书馆就能存储远超个人所能记住的信息。成千上万作者的思想和成就能被永远记住。有了互联网，这些信息就能被全世界的所有人随时查阅。

文化不仅仅是时尚、流行音乐、民主制度和宗教活动这些虚的东西，也包括厨具、家具、计算机和汽车这些实实在在的东西。所有这些都是因为复杂信息的交流才成为可能，包括指令、配方、规律、规则和

经验。这些东西的交流需要语言，而语言本身也是可以交流和学习的文化的一部分。

文化如何演变？

文化进化的思想出现的时间并不晚于动植物进化的思想，甚至还要早一些。达尔文在1859年发表了《物种起源》，在此之前，19世纪的许多学者，包括康德和斯宾塞，就论述了文化的进步。达尔文自己在《贝格尔号航海志》（1839）中也思考过“文化的进步”。达尔文后来又在《人类的由来》一书中扩充了这些思想。但达尔文没有尝试在文化进化和自然进化之间建立明确的联系。这个任务留给了后来的学者。理查德·道金斯在1976年提出了拟子一词，意指假想的文化单元在文化中扮演的角色就等同于基因在生物中扮演的角色，强调文化进化与生物进化的相似之处 [1] 。此后拟子的概念引发了许多讨论。在谷歌搜索拟子会返回三亿七千六百多万条记录。这个惊人的数字证明了这个词在西方文化中的影响力。这个词在被提出来后，又获得了新的意义，现在很多时候拟子一词指的是思想或概念在网络上的传播。

从科学的角度看，拟子的概念有两个缺点，一是很难用实验来检验，也无法像基因一样准确界定其物理结构。基因就是DNA分子，是核苷酸序列组成的能携带信息的特定化学结构（参见第4章）。基因有具体的作用；编码蛋白质或RNA分子，或者作为调控蛋白的结合点。而作为文化单元的拟子的概念则模糊得多。因此有批评意见认为这个思想太难琢磨，无法作为一个严肃的科学理论的基础。这也意味着拟子的概念不能直接适用群体遗传学家采用的那些成熟的方法。群体遗传学研究的是生物群体中基因的行为如何随时间变化。基于拟子建立的文化传播的数学理论的尝试缺乏群体遗传学那样的严格性和精确性。道金斯自己从未辩称拟子的思想是严格的科学理论，而是将其视为一个有用的修辞手法。

不过，更宽泛地认为文化是基于信息的共享，而不是将细胞中的信息使用和社会中的信息使用具体等同起来，也可以带来深刻的洞见。将进化视为计算的好处之一是即便缺乏对文化信息的物理结构的精确界定也能探讨文化演变的问题。有了信息的最小单位比特的概念就足够了。一个可以探讨的问题就是文化的演变是不是也像生物的演变一样是进

化。

文化进化与生物进化的一个明显区别是繁殖的概念。生物繁殖时，信息（以DNA序列的形式）从父代传给子代。复制的信息以化学结构的形式进入新的个体。类似的，身体产生的新抗体需要有具有新的DNA序列的新一代B细胞（第7章）。大脑中思想的进化涉及一系列神经元通信网络的形成，每一个都是大脑中不同的物理模式。对于文化，虽然可以将思想的学习视为一个思想在另一个人的思维中“繁殖”，但扩增比繁殖更符合发生的事情 [2] 。无论怎样描述，人与人之间思想的交流必然涉及某种形式的复制。

文化是基于共享的信息，个人参与文化的重要活动就是信息的传递和选择记忆。复杂引擎的计算的3个要素（图5.2）是：复制的信息体、不完美的复制（例如复制时以一定的概率产生适度的变化）以及选择确定哪些复制体会被再次复制。图11.1重新绘制了这个循环，可以看到这本书中给出的广义进化定义同样适用于人类文化的交流。在第8章我们看到复制过程中的变化不需要是完全随机的，虽然随机使得未来的结构有许多可能。关键在于要有不同的输出，从而为选择提供运作的基础。如果所有输出都是一样的，就不会有变化。这个框架有一个有趣的变化，许多选择规则似乎都是从文化内部产生的。这样的例子包括流行文化的时尚，市场选择产品，以及宗教信条随时间的改变。因此在这里是选择的规则经常是它们所运作的文化的产物。因此，大部分规则也是选择的对象。

我们都很熟悉口头交流的不可靠。小孩子“打电话”的游戏之所以有趣正是因为这种不可靠性。当参与者小心避免准确复述时游戏尤为有趣，但如果参与者完全改变消息就不好玩了。当参与者引入的变化很小时游戏最有趣。书写让信息的交流更为稳定，但即便是书写的文字，人们也发现在传播消息的时候很难不“修改”。修改就是变化，无论是否随机，任何变化都会为选择提供素材。

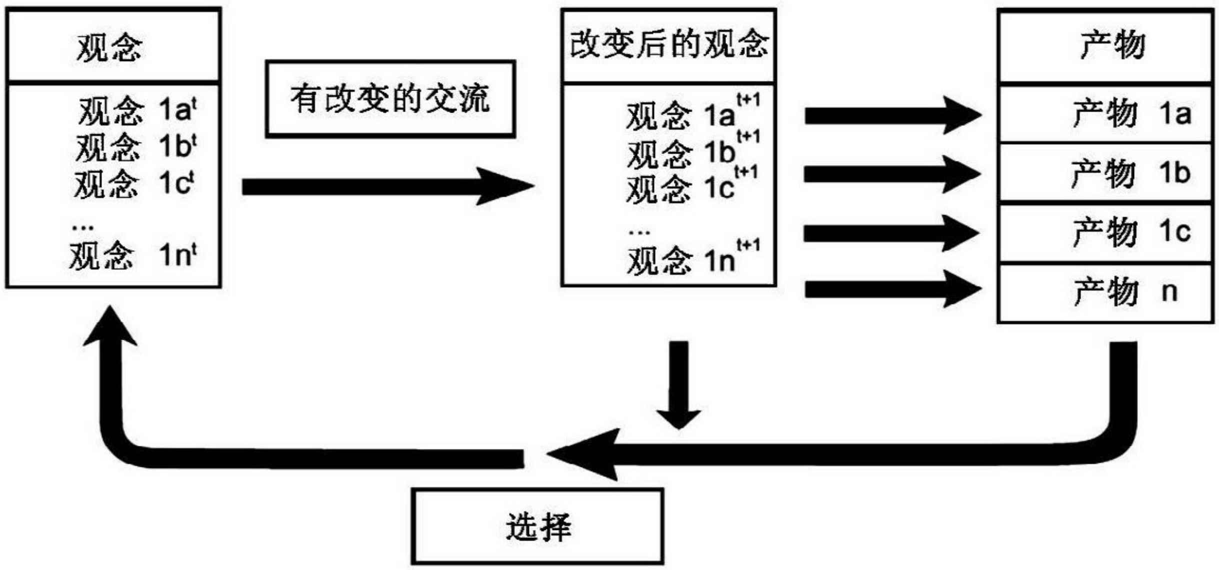


图11.1　重绘的图5.2的复杂引擎，表明了这个基本框架同样适用于人际交流。图中的输入和输出变成了“观念”，随机复制变成了“有改变的交流”。标为1a、1b的相同观念的各种版本同时在流传，每当观念被交流，无论是否改变，上标加1

但文化不仅仅是交流的信息；它是被广泛认可从而获得了一定稳定性的信息。图11.1没有明确展示个人是如何参与的。一般来说，共享观念的信息存在于人的大脑里（或是书和硬盘这样的存储介质里）。对于文化有贡献的信息必须有闭合的交流环。当观念在人际间传播，最终会以修改了的形式回到最初的交流者那里。只有回来的信息还类似于原来的信息才能认为信息具有了足够的稳定度，能在文化中更长久地留存。书写的信息由于更具持久性从而更容易达到这种状态。

为了说明复杂引擎是如何整合到文化的进化过程中，下面探讨3个具体的例子：科学进步、经济增长和宗教。还有很多例子也是一样，例如语言的进化、政治制度、技术进步、流行音乐，等等，只要是人类文化的组成部分，都会不断进化，只是这些进化是通过书籍等具有持久性的媒介实现的。

科学是怎样运作的？

科学知识对现代社会的影响无处不在，因此对于科学如何生产有用的信息已有许多研究。这方面最著名的两本书是卡尔·波普尔的《科学发现的逻辑》和托马斯·库恩的《科学革命的结构》。波普尔的书对科

学家如何认识自己所做的事情有很大影响，库恩的书则对人文学者的影响更大 [3] 。科学是复杂的文化活动，两位学者都为我们对这个抽象问题的认识给出了深刻的洞察。

根据波普尔以及大多数科学家的观点，科学本质上是谨慎地解决问题的过程。这个过程从科学家提出问题开始：一般是关于自然界某个他们不清楚的问题。然后是提出假说，如果假说为真就解决了所提出的问题。这两个步骤很平常，我们每个人每天在日常生活中都会进行很多次。让科学与其他人类活动不同的是，提出的假说要能被实验检验，否则就没多少意义（或价值）。对于自然世界很容易构想各种解释，比如在月球的背面有外星人或者森林里有精灵，但要想出能被检验的解释就不那么容易，而且经验表明，无论假说多么吸引人，其中绝大部分都是错的。这是一个显见事实的直接推论，即错误解释的数量总是要比正确解释的数量多得多。如果对某个事物的相信缺乏证据，则其正确的几率极小。只有假说被实际的自然现象证实，它们才具有可信度。大多数科学家的日常活动就是设计和进行各种试验，以对假说进行检验。这也就是为什么“科学神创论”让生物学家很烦心。倒不是说对世界的宗教解释就一定是错的，而是因为它们无法被检验。没有实验检验，解释就不是科学，对自然现象的非科学解释无法让人满意。当它们可以被检验时，往往都被发现是错的。

科学不那么明显的一面是，实验的目的不是证明而是证否！从原则上无法证否的解释也许会很有吸引力，但终究没有价值。当一个精心设计和实施的实验与假说一致时，并不能证明假说，而是增加其可信度。如果实验结果与假说相矛盾，则说明要么对实验的解释是错误的，要么假说需要被修正。

科学的另一个重要特征是实验及其结果必须发表出来供其他人参考和重复。个人的逻辑很不可靠。我们大家都喜爱自己的思想，看不到自己的偏见，科学家也是一样。只有与假说的关系不那么密切的人挑剔地检查了每个实验的设计、实施和解释之后，实验结果的有效性才会被采信。即使是这样也还会有隐藏的问题。当假说以及支持假说的实验被发表后，其他专家接纳后也还是会提出新的方法来对其进行检验。只有各种实验得到的结果与假说都一致时，假说才会被广泛接受。即使是这样，也没有哪个科学假说是地位永固的。人们不知道是否某个新的实验会发现之前没有发现的不一致性。牛顿的引力定律作为科学理论牢牢地

站立了两百年，直到20世纪的许多实验表明爱因斯坦的广义相对论能对引力现象提供比牛顿定律更完整的解释。结果发现牛顿提出的方程只是更普适地解释自然的方法的一个特例。这个特例在像地球这样的地方很有效，却无法描述在宇宙其他地方极强的引力场中发生的事情。也许还有更具普适性的认识引力的方法没有被发现。时间将揭示这一点。

对于刚开始学习科学的大学生来说很难理解的一个概念是，科学不是寻找终极真理，而是寻找“可行的最佳解释。”如果我们不知道所接受的科学定律和理论是否是真的，又怎么会以科学为基础产生了如此多的技术成就呢？答案是得到了证据支撑的科学理论在特定情形下已足以胜任，即使将来可能会发现解释同样现象的更完整的方法。定律在所限定的范围内足以支撑工程领域；如果你只是要建一座桥，牛顿引力定律在黑洞附近是否会失效并不重要。

生成假说，实验检验，然后修正假说，这个过程已经被证明是在由可能解释组成的庞大空间中搜索一个适用于我们的观测的解释的有效策略。这个过程与图11.1所描述的方法一致，因此符合复杂引擎的定义。本质上，这个过程是对知识与自然进行匹配。这个搜索的过程不是随机的，但是另一方面，没有哪个科学家知道自己提出的最新假说能否站得住脚。实验检验为思想的选择提供了客观的基础。没有这样的基础，就无法知道特定的解释是否与物理现实相一致。当假说通过了检验，被发表后其他科学家会用更多的实验对其进行检验，如果再次通过了，就再次发表。得到了很好支持的假说被写入教科书，从而被更广泛地传播和检验。传播和接纳在科学中扮演的角色就相当于繁殖在生物界扮演的角色。

关于科学如何运作的进化观点并不会贬损假说的创造以及实验的设计和解释这些智力活动的作用，而是对宇宙的复杂性的证明。宇宙太过复杂以至于谁也无法全部理解，这是老生常谈，但却是事实。

它的一个推论是谁也无法预测宇宙尚未被认识的那些方面在未来会如何表现。科学解释让我们可以预测自然被很好认识了的那些方面。从信息科学的角度看，发现新事物的途径就是搜索，而一旦要搜索的信息空间太大，无法用穷尽的方法搜索，而且你也不知道自己到底是要找什么的时候，图5.2和图11.1描绘的搜索引擎就是寻找符合你的目标的事物的最有效的方法。科学的方法就是在人类的社会环境中对这个技术的实

现[4]。

埃里克·拜因霍克用“演绎修补”描述人类思维在文化变异（产生新的假说，等等）中的作用[5]。当人们生成假说，无论是一个科学理论还是对汽车司机在开车时糟糕表现的解释，都不是凭空出现，而是基于他们已知的东西。对于科学，一个（好的）新假说必须是基于更成熟的理论并与之一致。有许多违反热力学定律的理论曾被提出来。由于热力学定律被以各种方法检验过很多次，因此要在我们的宇宙中的某个角落发现其被违反的可能性极低。与普通科学家相比，伟大的科学家的一个特点是他们对被其他人所发现和接受的东西有更好的理解。他们能调和所有已知的信息，因而能提出更好的假说，但即使是伟大的科学家也经常猜错。

被接受的科学理论和定律组成了一个信息体，其中涵盖了我们知道的关于宇宙的一切。本质上科学是收集、保存、修正和增添这个信息体的社会事业。这些信息以书面形式被记录，存留在活着的人的大脑中，科技期刊的文章中，以及文献综述和课本中；它是通过反复地生成变体（形成假说）然后选择（实验检验和选择保留）创造出来。归根结底，科学知识是许多科学家提出问题、猜测答案、检验猜测、修正猜测然后再次检验的学习算法的产物。图11.1形象地描述了这个方法。让科学与其他许多人类事业不同的是选择与自然一致的好思想的客观方法。通过数百年的努力，这个过程所累积的知识已相当惊人，并支撑了技术的不断发展，让人类能够以前所未有的程度控制我们的环境和命运。

经济会发展，但会进化吗？

在《财富的来源》一书中，埃里克·拜因霍克有力地阐释了经济的繁荣和衰退是一个涉及商业计划的改变、增长和选择的过程。简单地说，商业计划是通过制造和交易物品以赚取利润的策略。商业计划是商业活动和企业的指令。商业计划可以成文，也可以不成文。在传统经济理论中，除非交易双方都认为自己能从交易中获益，否则就不会有交易。因此交易不是零和博弈；交易越多，产生的价值就越多。

很久以前商业都是物物交换，一种东西换另一种东西，不涉及价格。在货币发明和被接受之后，大部分交易都是以钱易货。因此，制造

和买卖的物品获得了可以用钱衡量的价值。在经济理论中，价值的确立是通过买方愿意出价而卖方愿意接受出价。当某种物品被大量买卖，而交易的价格被大范围传播，则其“价值”就大致是许多次交易价格的平均。如果根据商业计划生产出来的物品没有人要，则计划就会被放弃，不能为卖家提供足够价值的商业计划也无法长期存在，因为卖家不会继续生产。许多计划被提出来，然后被市场选择淘汰。我们可以认为它们灭绝了。另一方面，好的商业计划能为买卖双方产生价值，从而被其他人模仿。模仿是商业计划在经济中传播（也可以说繁殖）的机制。通过演绎修补，人们不断“改进”现存的商业计划，市场则选择留下那些具有优势的，淘汰劣势的。公司用赚来的钱发薪水和买原材料；雇员和原材料供应商反过来又用他们获得的钱买其他商品。这样，钱就循环起来。购买和销售的物品越多，每个人的财富就越多。

商业计划不仅仅是贸易，也包括物品的生产。随着计划通过反复修正和测试不断改进，它们就会自然积累信息。结果使得改进的产品和服务具有了更大的价值，因为它们的生产效率变得更高（提高生产率），也具有了更多人们所需的特性。

拜因霍克阐释的财富产生机制的一个中心特征是合作提高了生产的效率 [6] 。现代公司很善于产生财富，因为它们很善于寻找合作的机会来生产产品和提供服务，以更低的成本更好地满足人们的需求。效率的增加可以认为是改善了不同目标之间的契合。在市场中总是存在能创造价值的策略和不能创造价值的策略。商业合作计划是利用可能机会的特定策略。丰田汽车公司采取的就是将许多人组织起来生产质优价廉的汽车吸引顾客的商业计划。购车者认为自己的购买是一笔“好买卖”，否则他们就不会买，而车的价格也能让公司挣到钱，否则他们也不会卖。各部门效率的提升使得车的价格降得更低。现在丰田已经成为世界上最成功的汽车公司之一，它的商业计划的许多方面也被其他汽车公司复制，以使他们自己能变得更成功。

如果你不相信合作行为在丰田的商业计划中扮演了重要角色，只需想一想如果你自己用零件组装一辆丰田凯美瑞要花多少时间和精力。然后再想一想你自己去收集原材料来制造零件所需的成本和精力。没有人能做到，更不要说还有采矿、冶炼、锻造成型、机械加工，还有生产线材，制造计算机芯片，以及用于绝缘、座椅和保险杠的各种塑料的成型。你还需要一个化学实验室提供化工原料生产轮胎、车漆和各种塑

料。这样生产汽车的成本很难估算，没有意义，就算有人有这么多钱也没这么多时间。如果每个人都自己造车，将没有人负担得起。早在两百多年前，亚当·斯密就在《国富论》中指出了这一点，他阐释了别针的制造被分为多个步骤，每个步骤由不同的人执行，从而使得别针厂的生产效率大为提高。

现代世界的经济中到处都有这样的例子，不同专业的人通过合作使得他们的产品价格远低于单独作战的人生产的同类产品。生产率的提高带来的低价使得销量增加，多赚的钱让雇员的收入更丰厚。获得的收入反过来又用于购买其他人（根据商业计划）组织起来高效低成本生产出来的商品。现代经济就是由大量这种组织活动相互关联形成的复杂贸易网络。这个系统很像第9章讨论的细胞生物化学/蛋白质组/基因调控网络，但有完全不一样的节点和边。这个网络结构是以商业计划为基础，商业计划是说明经济中各项商业活动细节的指令。

拜因霍克认为经济是由许多小的复杂适应系统（CAS）组成的复杂适应系统。他对CAS的定义是大量智能体根据规则交互组成的系统，在其中智能体具有处理信息的能力，并且根据信息调整各自的行为。在经济中，智能体既可以是个人也可以是组织。过去30年里，科学家对CAS已经进行了大量研究，并且发现了一些特征。首先，CAS的行为可以非常复杂，以至于不可能详细预测。这种系统经常会表现出涌现模式和行为，而不仅仅是各种输入和组成部分的简单加总。第二，这种系统各部分的智能体和外界输入的交互会倾向于产生大量短暂的稳态。第三，一个CAS经常很容易从一个稳态变到另一个稳态，而变化的规模通常遵循幂律分布。第4章讨论的沙堆崩塌就是幂律行为的简单例子。

而对于经济中的商业策略，营利性组织代表了暂时的稳态。由许多策略的互动组成的经济整体也表现出一定的稳定性。在这个系统中，总是有办法挣得到钱，也总是有更多办法挣不到钱。必然有一些商业会更为成功。真正好的办法也就是丹尼尔·丹尼特所说的“好点子”[7]。可行的好点子取决于当前的经济状况，机会不断变化。但总是存在一些好的点子，一旦被发现和实施就能挣到钱。问题是无法预先知道哪些策略是好点子哪些又不是。另外也无法知道一个好点子能多久有效。与不确定的环境进行交互的唯一途径是不断尝试新的思想并跟进有效的，也就是演绎修补的过程。成功的思想经常是通过对原有策略进行修正得到。

由于人们的需求不断变化，技术进步也使得新的事物不断出现，因此经济不断变化，商业计划的成功往往不会持续太久。任何商业计划一成不变的组织很快会发现其他组织能提供更多价值，不变的商业计划会失败。这正是复杂引擎意义上的进化。

从西方到全世界，经济表现出两个惊人的特征，前所未有的财富产生和前所未有的商业计划的复杂性。根据拜因霍克的观点，这些特征都关联在一起。就像前面讨论的，财富的产生是源自合作行为带来的效率提升。最好的策略组成好点子，一旦被实施会以有吸引力的价格生产出人们想要的东西。拜因霍克更进一步，认为财富不仅仅是钱；还代表了经济王国的某种秩序。他称之为契合秩序 [8] ——能与经济网络表现的力量和复杂度很好契合的秩序。根据这种思想，财富衡量的是经济的目的性结构。

为什么商业计划会倾向于变得越来越复杂呢？答案由两部分组成。首先，在越来越大的组织之间可以发现合作优势；虽然众所周知大规模人类组织会倾向于变得官僚、低效和不能响应环境变化，但还是会带来效益。自由市场所扮演的角色等同于环境在生物界扮演的角色：如果低效压制了规模的优势，其他小规模的商业计划就会胜出并迫使规模大的改变或消亡。复杂度增加的另一个原因是创新会产生出以前不存在的机会。在生物学中称为“生态位构建”。当新的机会被加入商业计划，它们就会变得更大更复杂。

因此进化经济理论用大规模带来的新机会和可行的好点子的效益解释了公司规模似乎不可避免的增长。毫无疑问数字革命及随之而来的通信、数据存储和信息分析效率的增加带来了几十年前不存在的新机会。一些公司必然会比其他公司更具优势。

宗教进化吗？

如果外星人访问地球研究人类，他们的报告中肯定会有一章专门讲述宗教在大部分人类社会中的重要性。也许他们会对这些宗教内容的差异之大感到好奇，尤其是神的角色和能力，虽然各宗教的信徒都认为只有自己的信仰才是正宗的。另一个有意思的特征是神从不直接显灵。凡人要想与神沟通，要么必须通过一个媒介（神父或巫师），要么就是信

心很强，无需超自然信息的证据。许多人深信超自然力量的存在对他们的日常生活有影响，但他们无法向不信的人证明这种存在。毫不奇怪科学和宗教经常会不一致。

除了神灵，所有宗教还有一个显著特征就是其中的人类组织结构。所有宗教都用规则和教条进行组织，并且有专人维护这些规则和教条。宗教也随时间变化。群体遗传学家大卫·威尔逊在《达尔文的大教堂》一书中令人信服地描绘了宗教进化的过程 [9] 。他的核心观点是新的宗教不断产生，能延续的宗教除了能给信徒带来精神上的好处，还能带来可观的经济和社会利益，并且这些宗教还具有会随时间变化的长期组织结构。我要补充的是宗教可以从信息的角度来认识（对于所有社会组织都是这样）。所有宗教的学说、传统和组织结构都能而且经常会被写下来。这些信息的功能与商业合作计划很相似。

任何宗教的成功或失败都取决于能不能吸引和留住信徒。不能吸引和留住信徒的宗教注定成为历史的注脚。罗马神殿的神之所以消亡是因为他们在与基督教堂争取信徒时失败了，（吉姆·琼斯创立的）人民圣殿教的结束是因为集体自杀的幸存者放弃了他们原来的信仰。

所有宗教都承诺一些无法证实的事情。原始社会的人们会遇到许多无法解释的事情，他们的宗教提供了各种机制（祭祀、仪式，等等），让他们在面对喜怒无常的神灵的率性时能获得心理安慰。农业和劳动分工的出现使得大型社会的出现成为可能，成功的宗教会许诺永生的希望。这也许反映了在艰难的社会环境中生活的人们的悲伤和绝望。现代经济让许多人脱离了赤贫状态，随着条件的改善，许多基督教派倾向于越来越强调生活的满意度。人们受那些能与他们的处境共鸣的信息吸引。就这样，宗教进化了——他们的商业计划发生了改变。不适应的则逐渐消亡。

威尔逊指出成功的宗教倾向于提高信徒的生活条件。所有宗教都有为信徒的组织合作行为提供支持的组织结构。这种合作经常能为参与者带来可观的经济和社会利益。群体生物学家深入研究了互利如何通过自然选择在动物中出现。困难之处在于自私的个体在面对利他者时具有明显的优势。但如果合作者同意集体惩罚自私行为这种优势就会消失。人类的合作需要合作者之间的信任和有效沟通。成功的宗教会有鼓励合作行为的规则。坚持所有信徒必须遵守规则使得信徒可以信任不熟悉的

人。这使得大规模的合作行为成为可能，在合作行为的世界中有许多好点子能为合作者提供可观的好处。成功的宗教总是有明确的手段促进"好"的行为。好处是通过加入合作社团，生活条件得到改善。心理承诺对于新的宗教很容易做到，但社团中合作的经济利益会对潜在的信众产生强烈的吸引力。所有能长期成功的宗教都利用了人类社会组织的这个根本现实。

威尔逊认为宗教的进化有两个层面。首先，他们相互竞争信徒。如果有机会，人类会在多个宗教中选择一个参与。因此，宗教必须表现出吸引力。在这个层面失败了的宗教会因为没人感兴趣而出局。由于新的宗教不断产生，总有选择的余地。除非某个宗教赢得了区域垄断并且组织制度的设置使得退出很难或很危险。进化的第二个层面是界定一种宗教及其运作方式的规则、教条和结构。需要对教徒具有长期的吸引力意味着所有宗教的细节都要不断修正；不断尝试新的特性。在这个层面上，宗教的进化与公司的进化没有什么区别。两者都要与其他组织竞争顾客，也都有不断进化的基于信息的企划计划。

无论是公司还是教堂的企划计划，进化都是通过组织中的个人行为进行小的增添或删改。然后这些改变被组织内部或组织之间的选择机制检验。通过这个过程，宗教的方方面面，甚至其秉持的最深的教条，在长时期里都会变得很不一样。这些变化可以很剧烈，就像耶稣基督从犹太教中出现，或约瑟夫·史密斯创立摩门教，也可以很缓慢，就像公理教会从定居在马萨诸塞的禁欲的清教徒演变成更宽容的联合基督教会。在所有这些例子中，现存思想和行为方式的信息内容被修正，并接受潜在信众市场的检验。

复杂引擎在人类社会中如何运作？

作用于信息体的复杂引擎需要有复制和修改信息的机制，还需要有对这些修正进行一致选择的环境。无论是将社会作为一个整体还是分开来审视大部分社会组织，都表现出了所有这些特征。争议在于信息的性质以及复制的模式，但并不是那么神秘。基本上，支撑社会组织的信息存留在人们的头脑中。形成记忆的神经元信息模式被印刷和数字化的信息增强，但最终都要归结到人们听到或读到并且记住的事物。因此，对社会进化的认识与对我们如何编码记忆观念和思想的认识密切关联。复

制的模式与我们在前面的章节中看到的不同。让社会进化得以发生的复制与人与人之间的信息交流有关。这种信息可以读写，或是表现为2维或3维视觉形式，但只有被记住了，才算被复制了。将书放在图书馆中并不足以将新的思想引入社会，得要有人去阅读和记住。

记忆信息的修正不是问题。人类记忆并不可靠，人类是典型的每次交流记忆都会篡改的叙事者。对于研究社会进化的人来说最大的困难也许不是变化如何发生，而是新的观念在交流中如何保持足够稳定的形态。

对于复杂引擎在社会层面的运作有一个必须提及的重要特点是一致的选择。选择标准可能改变，但如果变得太快，就无法累积任何东西。社会具有规范和道德感。认知科学家认为，我们的心智尝试通过叙事和思维模型的构建给世界赋予秩序。当这些构建被思维群体广泛共享时，他们就形成了牢固的世界观，在他们接受或拒绝遇到的新思想时为他们提供一致的基础。如果一个思想与内心的规则集不匹配，通常都不会长久。

这个简要的讨论应该已经解释清楚了，社会组织——无论是公司还是宗教——同橡树、细菌、免疫反应和思想一样，都是借助复杂引擎进化。所以这些最终都基于同样的计算策略。下面我们转到进化系统的一个常见现象，即复杂引擎通常能导致不断增长的复杂性。我们在现代经济中看到的一个例子就是公司不断增长的复杂性。难道复杂引擎具有某种内在的驱动力会让它的输出越来越复杂？还是所有进化系统都面临特殊的压力？

第12章　复杂性的进化

复杂性可以定义吗？

对进化会导致复杂性不断增加的认识深深根植于西方文化。这出于一些很明显的理由，但科学对此的认识则是另一回事，并且依然没有解决。一方面是令人炫目的科技进步，以及生命从简单化学混合物到具有自我意识的人类的发展。这些例子表明有不可抗拒的驱动力使得事物越来越复杂。另一方面则是许多生物谱系在很长时期内复杂性并没有明显增加，实验也表明进化算法可以被设计成既能增加复杂性也能减少复杂性（第6章曾讨论）。在数百万年里都没有表现出明显变化的生物谱系的两个例子是蟑螂和马蹄蟹。寄生虫谱系则在进化中明显表现出形态复杂性的减少。

这些例子都利用了同样的复杂引擎计算策略，但有不同的表现形式，处于不同的环境中，运作不同的选择规则。显然，细节很重要。大量关于进化算法的研究提供了令人信服的证据，运作在计算机中的复杂引擎并不会一边倒地驱动复杂性；决定长期趋势的是系统的细节。有两个显然很重要的细节是变异产生的机制和具体的选择规则。复杂引擎在每一代出现的变体中进行选择，如果选择的范围受限，将来的结果也就会相应受限。极端条件下，如果完全没有变化产生，引擎就会一直维持现状。如果变化来自有限的可能集，引擎就会穷尽所有可能，并停留在可能的最佳组合，但如果变化不受限制，探索就不会有尽头。很小的随机变化就能有效地实现这一点。

选择的规则也很重要。进化算法通常会计算输出的一些属性，并根据最好的成绩选择下一代。这类算法产生的输出的复杂性既可以增加也可以减少，取决于如何度量和选择。简而言之，复杂性可以选择增加或是减少。

生物中的选择是基于繁殖的成功，并不涉及计算。繁殖的成功取决于内部因素、与环境的互动以及机遇。进化生物学家在理解生物复杂性的来源时遇到挑战是在生物特征、环境和复杂性度量之间找到令人信服

的关联。如果复杂性能带来优势，进化当然会倾向于它，但难道复杂性就一定会带来优势？在社会的变革中也有类似的问题。更加复杂的社会组织比简单的组织具有哪些优势？就目前的认识水平来说，答案似乎依据文化的视角不同而不同。例如语言的复杂性就没有表现出必然的增加，而技术和经济的复杂性则至少在最近一段时期表现出了增长。抗体的产生则完全不同。当我们受到感染，产生的抗体会越来越适应抗原。复杂性似乎不是刻画这一点的合适方法，亲合力和选择性是更合适的度量。

在讨论中我是以一种非技术性的常识性的方式使用复杂性一词。我们直观上就能感觉到一些事物比另一些更复杂。但如果要深入探索进化与复杂性的关系，我们就需要仔细对这个词加以定义。如果我们认为有什么东西增加了，唯一能确证的方式就是进行度量。很显然人类要比变形虫复杂，苹果电脑也要比算盘复杂，而且不难找到符合这种观察的度量；但又应该如何比较苹果电脑与变形虫的复杂性，或是微生物群落与人类社会的复杂性呢？并不是说还没有发现对复杂性的度量方式。已经提出了很多方式，但还没有哪一种能抓住我们直觉认识到的复杂的全部方面。复杂性自身的意义本来就很复杂。

许多物理学家认为负熵是复杂性的合适度量。负熵是系统的熵与同一个系统处于平衡态时的熵（此时具有最大的可能熵）的差值。由于熵度量的是无序的程度，因此负熵就是对有序的度量。就此来说，一个信息体的负熵就是其携带的某种事物的“信息量”[1]。负熵定义的问题在于虽然有序和复杂性看似相似，却并不是同一回事。两个复杂性明显不同的系统可以具有同样的负熵。香农信息量与熵在数学上是等同的，而且就如我们在第1和第2章看到的，这种度量并没有抓住复杂对象的重要特征。用系统的最大熵减去实际熵并不能改变这一点。

复杂的事物必定由许多部分组成，并且各部分之间有许多相互作用，据此一些人提出应当区分“结构复杂性”与“功能复杂性”[2]。就好比一辆汽车与装着同一辆汽车拆解后的所有零件的大盒子。两者具有相同的结构复杂性，即由相同的部件组成，但部件之间的组合和作用方式不同。显然，组装好的机器在某方面比装着相同零件的盒子更复杂。因为功能正常的汽车能表现出一盒子零件不具有的行为。

结构复杂性的概念中包含了部件的数量、部件种类的数量以及结构

尺度。尺度很重要。将圆珠笔拆开可以发现包括墨水在内的8个明显的部件，但笔还含有大量的原子。那么，在分析笔的复杂性时是不是应当考虑所有原子的位置呢？如果从笔上拿掉一个原子有没有问题？一般来说，在确定宏观事物的结构复杂性时，必须说明感兴趣的尺度，否则就会陷入不相关的原子和亚原子细节。

功能复杂性包含行为、对行为的控制以及时间尺度。尺度同样很重要。一些事物发生得太快或太慢，人类无法察觉。这些通常可以忽略，不会有负面效应，完全取决于我们对什么感兴趣。弗朗西斯·海利戈恩就曾强调，复杂性的评估总是相对于观察者[3]。这是因为所有观察者都有一个天然的尺度，而在给定的背景下，只有特定的方面才是重要的。

研究复杂性的另一个方法是考虑其反面，简单性。简单事物的部件更少，也许只有一两种部件，比如二进制字串中的0和1。当我们比较具有相同数量部件的事物，简单的倾向于重复，复杂的则较少这样。结构的组分可以像晶体中的原子那样在1维、2维或3维上重复，也可以像声音的模式一样随时间重复。同样，重复也可以是简单的或是复杂的。

物理学家用对称分析重复。一种很简单的对称是镜像对称：事物的一半就像另一半的镜像。要描述镜像对称，只需描述其中一半，并说明另一半是这一半的镜像。可能的对称的种类是有限的。一些简单的情形包括2重平移（就好像复印）、3重旋转（三角形）、随时间重复（波）。对称对于理解复杂性很重要，因为事物越对称，所需的描述就越少。6重对称比2重对称更容易描述，因为只需要描述1/6的细节。振荡器一遍又一遍重复相同的模式，但要描述这种行为只需要描述一次基本的重复。对称也可以叠加；例如一个事物可以同时表现为2重和5重对称。晶体既有平移对称也有旋转对称。从这个角度看，最复杂的事物表现出的对称也最少。

沿着这条思路推到极端，意味着气体这样的组分随机排列的事物是最复杂的。但气体不具有结构，而结构通常与复杂性相关联。避免这个尴尬局面的办法是认识到随机对称性。气体的任何一部分都与其他部分在统计上是一样的，因此当考虑统计平均特性时，气体高度对称，从而很简单。总结这条思路，复杂事物具有复杂的对称。我不认为这个定义作用很大。

计算理论提供了另一种分析复杂性的方式。在计算理论中的对象是0/1序列，这种对象并不是凭空出现的，它们是计算的结果。计算视角有一个优势是数字输入和输出具有明显的天然尺度。计算机操作0/1符号时底层的电子的行为，或者印在纸上的0/1序列的墨水分子显然与计算的意义无关。重要的是0和1，而不是它们是由什么构成的。数字计算的另一个概念优势在于输入和输出仅由0和1两种简单组分组成，研究这种系统比研究由许多不同种类的复杂组分组成的系统更简单。

计算机科学家有没有刻画0/1对象复杂性的方法呢？我们已经见到了一些。最简单的是数0和1的数量。第1章介绍的香农信息是基于概率，一般用于度量对象整体，但如果对象属于某个整体，而整体各部分又具有相同的概率，香农信息也可以用来刻画单个对象。计算的输入也可以视为输出的描述。从这个角度看，输出一个对象的最短输入的长度就度量了其最短描述。一个对象的完整描述必定需要以某种方式包含其复杂性；因此最短描述的长度必定是其复杂性的一个特征。这就是第2章介绍的算法复杂性也即柯尔莫哥洛夫复杂性背后的思想。没有内部关联，组分随机排列的对象的算法信息量最大。从人类对复杂性的判断来说，这是一个让人失望的结论。

长度、香农信息和算法信息都没有抓住对象复杂性的一个方面，就是从短输入计算输出对象的难度。复杂对象由许多部分组成，并且部分之间有许多关联；因此如果一个复杂对象是由短的输入计算得到，则组分及其关系就必须在计算过程中创造出来。一个有吸引力的思想是，创造的对象越复杂，所需的计算量也越多。

这自然引出一个问题："计算量"到底是什么？在计算理论中，关键资源是时间和空间。空间是计算过程所需的内存（存储比特值的位置），时间是操作数据（在各种状态之间变化）以及在内存中存储和提取数据所需的指令周期数量。通常时间和空间（即内存）不可兼得。如果内存有限，计算也许仍然可以进行，但需要更多时间。反过来，经常可以通过增加内存用量来减少计算所需的时间。不过读写内存还是需要时间，因此即便内存空间不受限制，计算输出一个复杂的对象可能仍需花费很长时间。

这种复杂性定义所需的短输入确保了输出的细节没有简单罗列在输入中。通常在计算机科学中需要用短输入长时间计算出来的结构被认为

是“深的”，可以用短输入很快计算出的结构则被认为是“浅的”。深度反映了对象中存在的内部关联的复杂性。关联越复杂，从没有包含它们的输入计算出包含它们的输出所需的时间就越长。正是从这个意义上深度提供了对象复杂性的一个特殊度量[4]。

计算理论的一个基本贡献是从数学上证明了存在可以想象，但不能在实践中用比其本身更短的输入计算出的对象。不能的原因有两个。一是所需的计算可能极长，以至于虽然可以写出生成特定长对象的短算法，所需的时间却要超过宇宙的存在时间。另外在考虑所有可能的对象时，长的要比短的多得多。这很容易理解。1比特长的对象只有2个（0和1），2比特长的对象有4个（00、01、10和11），3比特长的对象有8个，4比特长的对象有16个，以此类推。特定长度的字串数量随着长度增加呈指数增长（2^1，2^2，2^3，2^4，……）。在这个数字序列中，最后一项是前面所有项的总和。因此，可能的长输出的数量总是多于可能的短输入的数量，因此大部分长对象必然只能用至少等长的输入计算得到。

要计算任何特定的事物我们必须从适当的输入开始。问题是：如果能生成所期望输出的最短输入很长，如何才能找到这个输入？通常，在计算机科学中，要寻找某个东西必须进行搜索。可以设计程序搜索可能生成期望输出的输入。如果这样的搜索从很短的输入开始，逐项检验越来越长的输入，则需要检验的输入的数量会随输入长度呈指数增长。这意味着随着检验的输入长度增加，搜索会变慢。如果想逐项检验长输入，所需的搜索时间会很快超过宇宙的存在时间。这是一个坏消息，意味着长输入无法逐项检验；而且穷尽搜索还会引出另一个更基本的问题。一些输入会使得计算永不停止；计算机会一直运行下去。也许一些输入在计算很久后最终的确会产生输出。不幸的是，无法知道任意的一个正在进行的计算最终是会停下来，输出重要的结果，还是会一直计算下去直到永远。这就是所谓的停机问题，意思是即使搜索的时间不受限制，也不可能通过对长输入进行穷尽搜索找出所有能产生预定输出的输入。

总而言之，有一些问题存在最佳答案，但没办法找到。这似乎很怪异，但是可以证明。幸运的是，情形并没有看起来那样糟，因为总是有可能找到能很好完成任务的次优输入生成所期望的输出。这也就是为什么能写出可用的程序。

用深度（以及第2章介绍的算法信息量）刻画对象的一个主要问题是大部分度量这些属性的尝试都受困于停机问题。除了极简单的情形，人们永远无法确定特定输出的最短输入或最短计算能否被找到。在大部分情形中这些度量都只是近似。不过深度和算法信息量的概念对于理解复杂性还是很有用。

生物学家如何定义复杂性？

在生物学中，复杂一词有多种用途。有时候它指的是地球上生命形式的多样性。在这种用法中，随着生物多样性的增加，生命变得更加复杂。这种用法在生态学中很常见，生态系统的复杂性取决于物种数量以及物种之间的相互影响。单个生物的复杂性则有不同的意义。人类是不是比鱼更复杂？应该怎样证明这一点？

对生命史有基本认识的人都会认同复杂性随时间增长。这可以从第一种意义的复杂性——多样性——得到证明。图12.1展示了化石记录中海洋多细胞生物种类数量的变化，不规则，但在统计意义上随时间增长[5]。第二种意义的复杂性——个体复杂性——也有证据。毕竟很少有人怀疑人类要比细菌复杂。

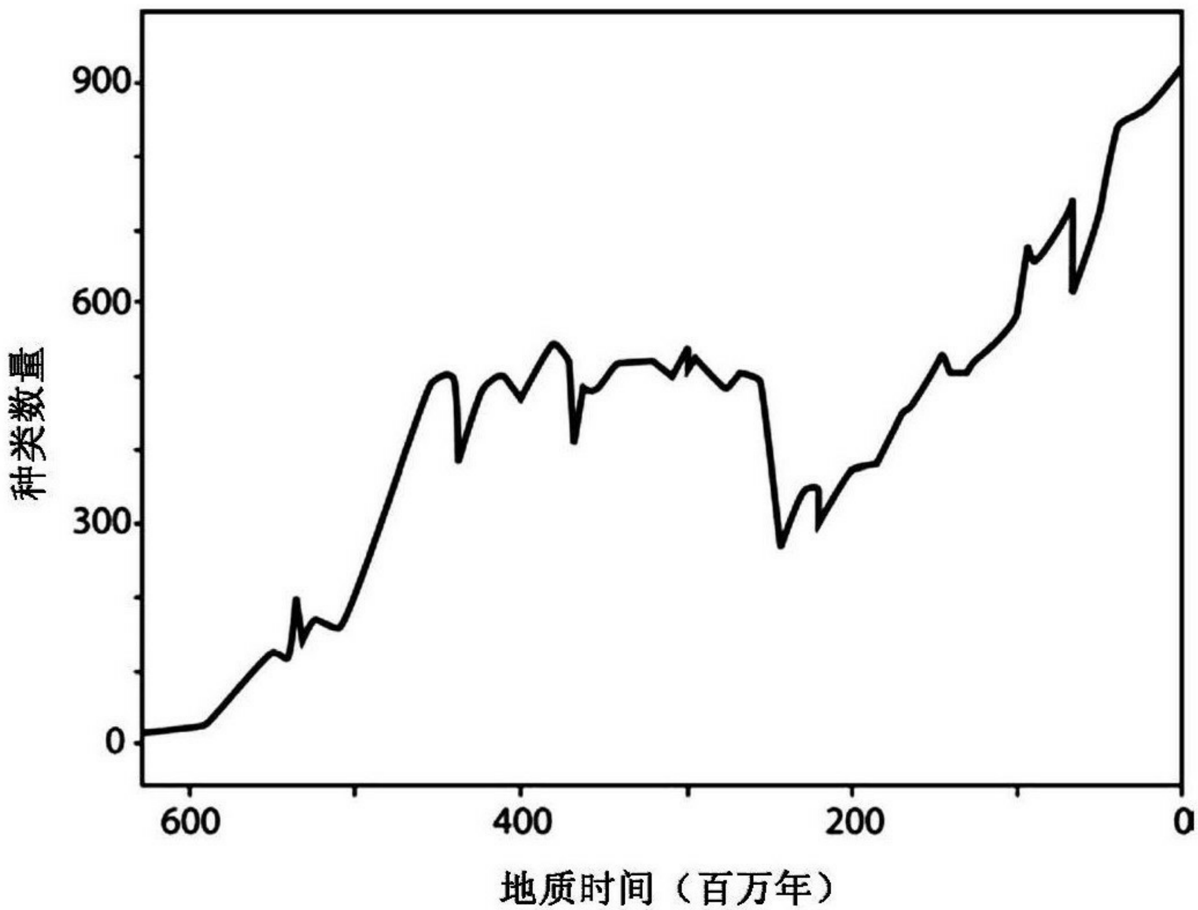


图12.1　海洋多细胞生物种类数量随地质时间的变化图。最初（最左边）的种类数量很少。现在（最右边）的种类数量是最多的（参见注释5）

当没有被完善定义的个体复杂性的概念与“先进”或“进步”的观念联系到一起时，就会产生争议。直到20世纪，西方文化中都普遍认为人类是逐渐进步的宇宙中的最高成就。这个思想至少可以追溯到亚里士多德的scala naturae，“自然的阶梯”，万事万物的排序，从最底层的岩石到接近顶层仅次于神的人。这个思想被融入了中世纪欧洲教会的信仰，在人类之上依次是天使、大天使和上帝。这个观念在西方思想中影响深远。一些人接受达尔文进化论，然后又错误地认为生命的进化是在亚里士多德的阶梯上攀升。很多人经常忽略了现代细菌进化的时间同人类一样长，并且如果以代来计算的话，细菌的进化史远比我们要久。根据一些指标，比如总数、总重量和适应各种环境的能力，细菌在总体上比我们更成功。我们很大很聪明，但谁又能说大和聪明就是判断进步的最佳方式呢[6]？

自《物种起源》于1859年发表后，达尔文的理论很快就与进步的观念以及进化具有目标方向性的论断纠缠不清。当然，目标就是人类，尤其是北欧人类。一般来说，在试图判断进化的进步性时，难免会与人类中心论牵扯不清。对我们来说，在评价其他生物时很难不认为它们比我们低等。在排序时这种无意识偏见会潜入进来。结果，经过这样的“科学研究”后，人类到达了顶点——正与我们在研究之前就已经知道的相符！很难做到价值中立。

热力学和/或信息论最有希望为进化产生的复杂性提供客观的度量。如果包含信息的结构是线性的，长度就能度量总量，但没有考虑重复。香农信息和负熵为累积的非随机性提供了可能的度量，但它们没有考虑计算的困难程度。算法信息考虑了重复，但在对象为随机时有最大值。这些似乎都不是我们想要的。作为从简单输入计算某物的困难程度，深度提供了另一种选择，但对于非数字对象我们就不知道该如何计算了。还有其他一些度量，但就目前来说，计算机科学还没有为生物复杂性找到合适的可计算度量。这意味着对生物学中复杂性的讨论缺乏坚实的理论基础。

尽管如此，无论我们怎样度量，人类还是明显要比单个细菌复杂得多。这很容易证明。我们由许多部分组成，并且各部分之间有许多互动。化石记录虽然很不完整，但还是很清楚地表明在很长时间里都只存在单细胞生物；很久以后才出现了海藻和奇怪的类似植物的动物。又过了很久，类似蠕虫的动物才出现在海里。再后来才出现了原始鱼类。只在相对很近的化石记录里动植物才来到陆地上并变得更加复杂。

陆生植物和陆生脊椎动物的化石记录都明显表现出更加适应陆地生活的形态。更加适应的标志是各种创新和应对干燥环境下生活的器官。最早的植物出现在大约4.5亿年前，类似现在的苔藓，即使到现在也还需要依靠雨水输送精子让卵子受精。后来出现了类似现代石松、马尾和蕨类的大型无籽植物。它们后来又被有籽植物——现代针叶树的祖先——取代。再后来，大约1.2亿年前，原始有籽植物又被有花植物取代。有花植物具有精巧的授粉策略和包裹在特殊结构中的繁殖器官子房，在其成熟后能提供散播种子的有力武器。蒲公英的绒毛可以借助风传播种子，葡萄和李子的肉质组织可以引诱动物（包括我们）帮助传播无法消化的种子，就是两个典型的例子。

在这个具有详细记录的宏大历史中，每一个大的“进步”都体现为新的植物特性。其中包括树叶上可以控制水汽蒸发的气孔，让植物可以长高的管道系统，以及将雌性繁殖组织包裹起来的胚珠，可以储存营养从而使得种子可以远离母体发芽。后来还出现了包裹胚珠的植物子房，可以更好地传播种子，蕨类等原始陆生植物的雄性繁殖器官也有巨大的变化，出现了能产生活动精子的配子体。在有花植物中，被称为花粉的雄性配子体是非光合的，很小，可以随风散播。无论从哪方面来看现代有花植物都要比苔类植物更复杂。

脊椎动物的化石记录也有同样的巨变，从鱼类到两栖动物、到爬行动物，再到最后出现的鸟类和哺乳动物，表现了类似的发展。

现代哺乳动物经历了从产蛋生物（单孔目）到胚胎哺乳动物（有袋类）到胎盘哺乳动物的发展，从而可以在母体内养育保护胚胎和胎儿。在灵长类动物的谱系中，最早的化石形态是生活在5500万年前的类似树鼩的小动物。然后是类似狐猴的树居动物，然后是猴子、猩猩，在大约700万年前出现了智人，我们的直系祖先。最古老的具有完整的现代面貌的人类骨骼只有大约20万年的历史，与35亿年的生命史比起来要短很多。

很难不认为这个进化顺序是进化持续进步的证据。但这一往直前的进步却伴随着原始生命形态的持续存在。细菌依然伴随着我们，还有海藻、两栖动物，以及树鼩、狐猴、猴子和猩猩。化石记录中的物种已经消失了，但经常有一些现代物种很类似其远古祖先。当然我们无法仔细检查远古生物的细节，以确定它们在分子和细胞层面上是否更简单。但是进化在一些谱系上前进的同时似乎还是留下了很多物种在数百万年甚至上亿年里都没有大的变化。

有许多线索已经让一些进化生物学家认为，进步的表象只是幻觉。首先，如前所述，许多谱系在时间的长河里基本没什么变化。根据化石记录，许多生物数亿年来都没怎么变化，蓝藻（类似细胞的光合细菌）、马蹄蟹和蟑螂只是其中一小部分，其中蓝藻甚至已经出现了十亿年。任何认为复杂性的增加具有内在驱动力的理论都必须顾及这一点。另外也有其他解释。最容易理解的是斯蒂芬·古尔德提出的著名的“左墙效应”[7]。

左墙效应是基于随机变化的统计论证。假设我们可以用数字刻画生物的复杂性，并将许多生物的数字画在同一幅图上，左边的复杂性低，右边的复杂性高（图12.2）。

无论复杂性怎样定义，必定有某个下限，低于下限没有生物能够生存。将这个下限称为“左墙”。当某个物种发生随机突变时，产生的后代比父代的复杂性要么高一些，要么低一些（或者一样）。图12.2 A给出了复杂性为A的父代生物及两个突变后代A′和A″。不可避免地，它们会落在图中父代的左右（或相同位置）。这样最初的父代位置的附近就会逐渐聚集越来越多的后代。所有后代的均值仍然保持不变，但极值会往左右两侧散播得越来越远。

再来考虑生命的起源（图12.2B）。在最初的生命刚出现时，生物必定非常简单，如果画在图上会离左墙很近。当后代在左右散开，右边的将会繁盛，而越过了左墙的那些则会消亡（复杂性不够）。图B中的A″就是这样的后代。如果制作一个影片展示经过许多代后各种生物的位置，（代表物种的）点将随机围绕着最初的父代。但由于最左边的后代（越过了左墙）会被淘汰，长期效应就是点逐渐往右边扩散（图C），如果只度量最右边的生物，就会发现复杂性逐渐增加。这正是化石记录所体现的！

左墙效应也称为扩散模型。还有类似的一些思想，丹尼尔·麦克谢伊将它们统称为产生增长的复杂性的无驱动或被动机制 [8] 。当被动机制起作用时，观察到的增加是随机变化导致的。这种模型并不认为有什么特性或谱系的复杂性会随时间增长。它们只是说总有一些生物要更复杂一些，如果被度量复杂性的恰好是那些在某个时候最复杂的，则得出的复杂性必然会随时间增长。

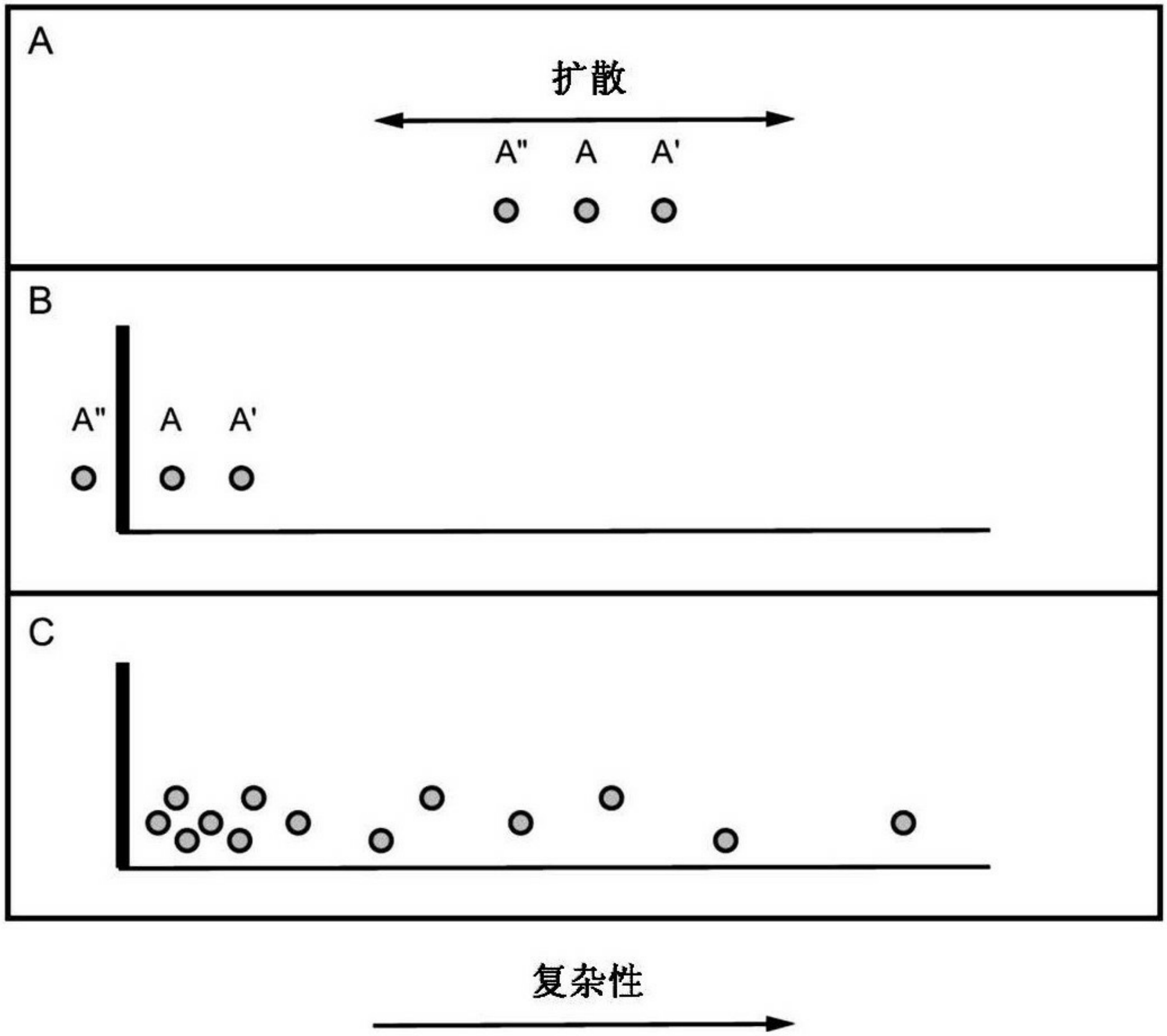


图12.2　左墙效应。复杂性低的生物画在左边，复杂性高的生物在右边。图A：点代表3种生物，位置由复杂性决定。A′和A″是A的后代。图B展示了生命史早期的父代和后代。生物A″越过了左墙，从而消亡。图C表示长时期后向右的扩展

另一个不同的批评意见则认为，已经证明很难找到与进化进步的主观认识相符的生物复杂性数值度量。如果具有一些分子生物学的知识，不难看到细胞DNA的数量给出了一种很直接的生物复杂性的度量。在20世纪70年代，当精确估计这个量的技术出现以后，很快就发现生物细胞DNA的数量与其表面的“进化进步性”之间没有直接关联。这就是著名的C值悖论。图12.3展示了各种生物单体DNA的数量级[9]。

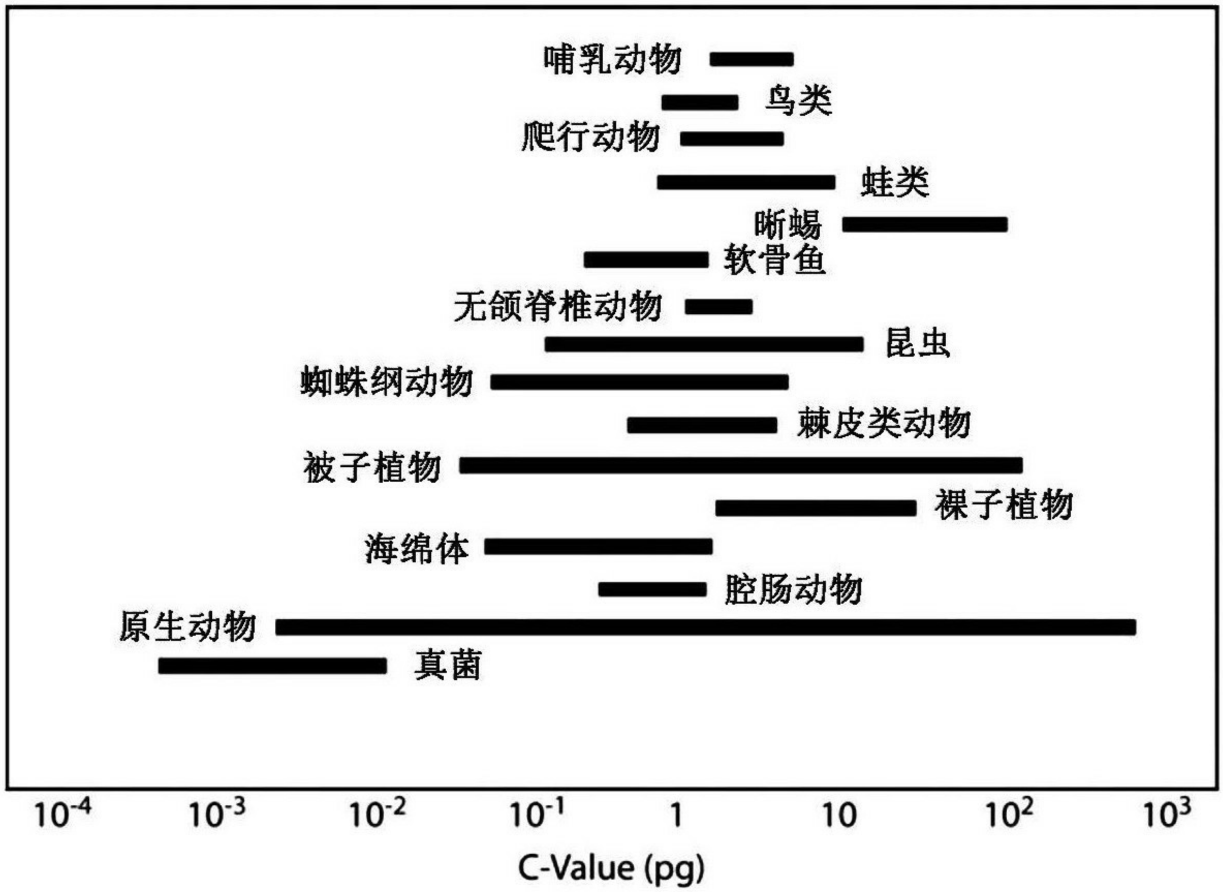


图12.3　单体细胞DNA数量（C值）图。大量生物都与进化“进步”的观念不符。长条表明物种的广泛多样。C值悖论指的是“进步性”与DNA数量之间没有明显关联。基于T.R Gregory, Paleobiology 30（2004）：179—202中的一幅图

虽然一些细菌（图12.3中的真菌）的DNA很少，一些原生动物的DNA却是最多的，一些植物和两栖动物的DNA甚至是哺乳动物的30倍。人类与其他哺乳动物的DNA的量大致相同，鸟类的更少，蜥蜴的则更多。最近35年来对DNA的研究已经取得了长足进展，首次发现了细胞中许多DNA并不编码蛋白质，最近又直接测量出了生物DNA中基因的数量。人类基因组计划一个最惊人的发现是人类的基因数量与线虫和墙水芹（属于芥菜系的一种简单开花植物）一样多。真菌和细菌的基因的确比人类少，但许多生物都比人类多。

与此相应，主要动物种群在化石记录中出现的时间与这些种群的现代物种的细胞类型数量之间则发现了惊人的相关性。图12.4展示了这种关联。

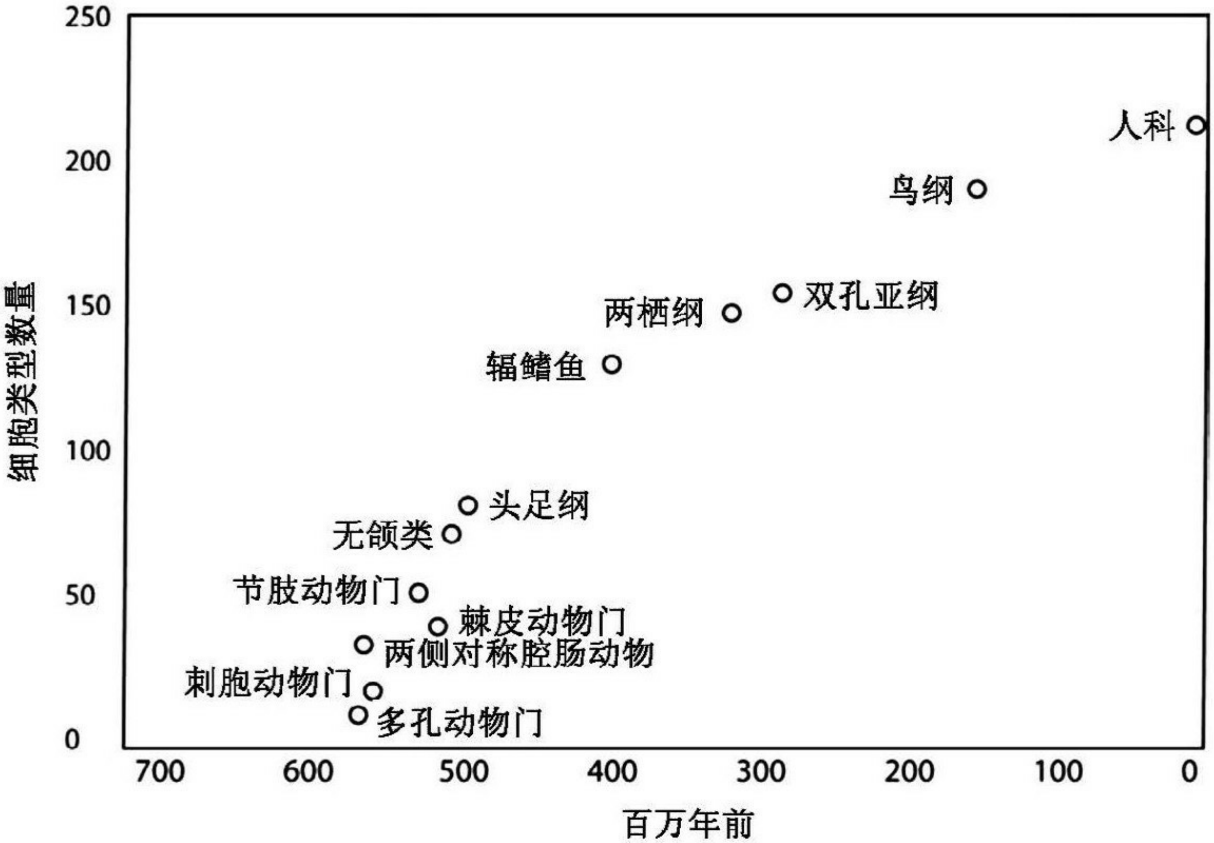


图12.4　一些后生动物的原始物种的细胞类型数量估计值与各类属的起源时间估计值的关系。经瓦伦蒂诺·凯伊许可复制：多孔动物门是海绵；刺胞动物门是珊瑚、水螅和水母；棘皮动物门包括海星和海胆；节肢动物门是昆虫和甲壳类动物；两栖纲是蛙和蜥蜴；鸟纲是鸟类；人科是人类[10]

批评意见认为细胞类型的分类有主观性，而且人类组织学的研究要比无脊椎动物组织学深入得多。然而，图12.4中的趋势不可能偏颇到完全是人为的。为什么用细胞类型数量度量的复杂性有稳步增长，基因数量度量的复杂性却没有呢？没有人知道确切的答案，但还是有一些调和这些发现的思想。

第9章介绍了基因和蛋白质相互作用网络对细胞的影响。节点一样多的网络，节点相互作用的复杂性可以差别很大。也许基因数量只能告诉我们节点的数量，而细胞类型则是对相互作用网络的复杂性的一个大致度量。背后的思想是各种类型的细胞的底层网络是一样的，只是状态不同。许多复杂网络都有大量的稳定状态，每种细胞类型体现了网络的一种特定的稳态。因此，观察到的稳态（细胞类型的数量）越多，底层的网络从而生物也就越复杂[11]。

最近还发现人类DNA的一大部分都是用于基因调控，细菌也是一样。到目前为止，还没有被广泛认可的可以识别和度量调控DNA信息量的方法。因此，我们还无法对其进行绘图分析。而且，第9章描述的生物化学、基因、蛋白质相互作用网络也远不只是基因调控。我们感兴趣的是整个网络的复杂性。生物网络复杂性的量化可能只有在这种网络可以用计算机精确模拟时才有可能实现。这在短期内还无法实现。

丹尼尔·麦克谢伊整理的一些例子清楚表明特定种系的确出现了复杂性的增加。这种例子很难找到。较完整的例子包括水生节肢动物在5亿年里的腿的类型数量，腹足动物（蜗牛）的壳在数亿年里的螺旋数的增长，名为菊石的已灭绝的软体动物的壳的连缝腔室的分形维度在3亿年里的统计增长。复杂性减少的一个明确例子则是从陆生祖先进化而来的鲸的脊椎的可变性。在这个例子中，似乎水中生活对脊柱结构的要求更简单，与其陆生祖先相比，进化显然更亲睐脊椎更均匀的鲸[12]。种系复杂性减少的例子有许多是向寄生性演变。寄生性生活形态的进化经常伴随着明显的结构复杂性减少。这可能是因为寄生环境要比自由生活动物的环境更简单。

研究复杂性问题的另一个途径是从复杂引擎开始，然后问通过研究背后的计算，我们能从复杂性的产生中学到什么。

最优化与协同进化有何区别？

在分析基于复杂引擎的活动的广泛谱系时，我们会遇到两种不同的情形。第一种是环境和选择标准不会随时间变化。在其中的进化实体会越来越适应环境和选择标准。有时候会找到最好的可能答案或匹配；有时候只是接近最好，但最佳答案至少在理论上是存在的。第二种情形是环境会不断变化；因此“最好的”答案只是暂时的。这一类中最有趣的系统是各个实体的环境本身就是由其他进化的实体构成。实体进化使得其他实体的选择规则也在改变。为了方便起见有时候将第一种情形（朝固定的最优解前进）称为最优化，将进化实体之间存在竞争的情形称为协同进化。两种情形都属于复杂引擎的范畴。进化系统只有在协同进化时才能发挥最大的创新潜力。

第6章讨论的将任意01序列转化为全1序列的最多1问题就是最优

化。虽然算法每次运行时搜索答案的路径不同，但最终的答案都是一样的。如果问题变得更复杂，最佳答案是什么可能不知道；但在理论上还是存在最佳答案，如果有合适的进化算法就有可能最终找到它。在运行了一段时间之后，如果在找到最佳答案之前优化终止了，很有可能找到的是一个较好的答案，如果运行更长时间，又有可能找到更好的。免疫系统应对感染就是最优化，最佳答案是什么不知道，但经过数天后产生出的结果往往已经足以保护身体免受进一步感染。身体需要这么长时间才能进化出能与特定外来抗原紧密结合的抗体。HIV病毒之所以让免疫系统难以应对就是因为这种病毒进化的速度不低于免疫系统的响应速度；因此身体产生的抗体总是脱靶。利用了复杂引擎的算法往往能解决复杂的最优化问题，但并不总是这样。朝最优解前进的速度取决于问题本身和所采用算法的细节。通常无法用进化方法解决的问题往往是那些最优解在解空间的地貌中没有被其他较好答案围绕的问题。这种情形很容易人为创造出来，但在自然界很罕见。

最优化的一个普遍特性是无法从整体上判断结果是趋向于更复杂还是更简单。结果的复杂性取决于最优解是复杂还是简单。而最优解是复杂还是简单取决于问题本身。

回到生命进化的问题，在自然界中选择往往是基于与其他生物的互动。换句话说选择是协同进化。在这种情形中，随着生物的进化，选择也不断变化，不存在永远的最优解。在计算机中可以设计出类似的进化算法，实体像生物一样受到其他实体挑战。甚至一个突变或随机变化就会改变整体的动力学。一个实体的变化会导致其他实体的选择动力学改变。在这种情形中会形成不断的挑战—响应—挑战循环。当其他实体改变，每个实体所面临的适应性地貌都会不断变化。在这个缓慢变化但最终不可预测的世界里，一时的优势可能过了一段时间就变成了劣势，没有哪个创新能确保长期的成功。人们可能会认为，在这样的系统中复杂性增加并不一定是好事，因为与复杂性相关的大量互动不大可能都从中获得好处，而且复杂性的具体特征也不会一直都带来好处。但结果并不是这么回事。

计算机进化算法带来的一个好处是现在可以进行以前只有活的生物才能进行的实验。在一个计算机实验中，实体相互竞争许多代，同时定期保存群体样本，比如说每1000代保存一次。然后将不同时期进化出的群体放到一起竞争。在其中第1000代可以与第10000代竞争。在生物界

这就好比将三叶虫保存下来放到现代海洋中去，或者将现代鱼类放到寒武纪的海洋中去。

通过这一类实验发现，进化了较长时间的群体一般比进化史更短的群体更具竞争优势。当两个群体是取自两次不同的实验，两者没有共同的历史时仍然是这样。长期的竞争似乎更青睐能应对新形势的策略。圭尔夫大学的丹·阿什洛克称这种现象为整体适应。他收集了6个例子，均采取很不一样的算法和数据结构，除了很少的代变化严重受限的情形之外，他还没有发现不表现出这种效应的系统[13]。

其中一个例子由丹的两个学生伊丽莎白·布兰肯希普和乔纳森·甘拉德完成，他们写了一个程序，在其中有两个虚拟的机器人在虚拟的12×12的格板上涂绘颜料。每个涂绘机器人采用不同颜色的涂料，通过涂绘颜料占据格子。机器人轮流动作，每次涂绘一个方格。机器人能够探测其周围8个格子的颜色。涂绘机器人只允许三种动作，右转、左转或前进。板子总共144格，每次游戏机器人各动作288次，足以将整个板子涂两遍。288次动作结束后占据格子多的机器人获胜。游戏中依次涂所有格子的机器人可能会输，因为对手只需跟在它后面涂就可以了。涂绘机器人的动作由被称为控制器的软件模块决定。控制器将周围方块的颜色作为输入（没有被绘制、对手的颜色或自己的颜色）并决定下一步动作。实验中机器人群体（实际就是控制器）进化许多代，每一代在机器人之间进行一定数量的竞赛。

在给出的实验中，有200个控制器（涂绘机器人），每一代每个机器人与4个随机选择的对手竞赛5次。得分最高的100个控制器被选出进行复制，后代产生突变并配对重组。重组是随机选取相同的位置将2个控制器（编码字串）截断，然后左右互配生成新的控制器。然后将重组和突变的控制器与100个父代一起放入游戏池进行下一轮（代）竞争。每一轮都用100个成绩最好的控制器的后代替换100个得分最低的控制器，最好的100个则直接进入下一轮。

通过反复竞争，结果发现如果最初的控制器是随机给定，则平均得分（机器人涂绘的方块数）在最初20到30代会迅速增长，随后逐渐稳定下来。在大约50代后，涂绘的方块一般是60个左右（总共144个），此后就不再一直增长，再过数千代也是这样。在最初阶段，完全无法胜任的机器人被去除，存活下来的策略则各不相同。288次动作足以让一个

简单重复涂绘策略的机器人将全部144个格子涂绘2遍；因此60分的平均分表明成功的机器人将大部分时间花在覆盖对手涂好的方块！用生物学打比方就是某个生物将大部分时间用于阻挠竞争对手而不是寻找食物。

这个实验最有趣的是将第500代的机器人与第5000代的机器人放到一起竞争。在99%的情形中都是进化程度更高的机器人胜出。虽然两代机器人在与同代机器人竞争时成绩都是60分左右，结果仍是如此。从统计上来说，纯属偶然观察到这种现象的机会还不到千万分之一（10^{-7}）。

其他很不一样的系统也得到了类似的结果，这表明进化更久的群体在某方面更加复杂。在涂绘机器人的例子中，决定实体行为的指令（控制器）的长度并不增加；因此如果以指令长度作为复杂性度量，并不会增长。不同的是发现了能成功应对不同环境的策略。适应范围更广泛的策略似乎能更有效地应对以前从未遇到过的群体环境。进化更久的指令在某方面比早期的指令更具适应性。分数没有增加是因为竞争更难了。可以说群体中的所有成员都变得更聪明了。类似的现象在生物群体中应当也同样存在。

资源重要吗？

一般来说，复杂性有相关的成本。复杂性的概念，除非定义为随机性，否则都意味着结构的存在；而结构的创造不是免费的。分析复杂性的成本和收益的一个有力而普遍的方式是将复杂环境下实体的成功与它们维持控制的能力关联起来。这其中的逻辑从生物学的角度很容易理解，如果个体的内环境不能维持在一定的限度内就会死亡。

控制作为工程领域的一个正式分支至少可以追溯到蒸汽机的发明。1956年罗斯·阿什比提出了必要多样性法则。这个法则说的是“要实现控制，控制系统能够执行的行为的多样性必需不低于需要补偿的环境扰动的多样性” [14] 。这个原理在电气和机械控制系统中有广泛应用，被用于例如加热和空调系统中或火箭姿态控制中。阿什比对其中的生物学意义有深刻认识。

用复杂性的术语说，阿什比的必要多样性法则说的是为了维持内部

的稳定性，控制器必须至少同其所抗衡的环境一样复杂。在生物学中，这个法则意味着进化会青睐于那些最善于维持控制其内部和外部环境的生物。不是地球环境的所有方面都与所有生物有关。对于任何生物来说，重要的环境复杂性取决于其生活方式和范围。对土壤细菌重要的不一定对老虎重要。由于环境复杂性如此多面，生物又各有特点，因此也就没有哪种通用的度量方法能被科学界普遍接受。

阿什比法则的一个逻辑推论是，当相互竞争的实体具有相似的复杂性水平时，任何生物（或控制器）都基本不可能实现对系统的全面控制。这是因为其他竞争实体的复杂性之和必然大于单个实体的复杂性。由此导致的控制能力的不足理应带来持续的选择压力，使得单个实体的复杂性越来越大。复杂性增加带来的优势和相应的成本之间的内在冲突必然会使得成本高效并且善于应对多变环境的特征具有特殊优势。

当环境不可预测，如果其他条件一样，受青睐的实体将会倾向于进化出能在多种条件下都能成功的结构和策略。当面对着不可预测的环境，能力越多的实体成功的概率也越大。

但能力并不能无限增长，因为任何能力都有相应的成本。能力增加成本也会增加。经过长期进化的实体必然反映了成本与收益的平衡。在物理世界，总是涉及4类成本，分别是能量、熵、时间和空间。

其中能量是最共通的。你的汽车没有汽油就开不动。我们的身体没有食物就无法运转。总的来说，任何物理设备都需要能量。对于生物来说，能量成本是显而易见的：运动，跨细胞膜输送分子，创造新的结构，修复旧的结构，都需要能量。对能量的需求是动物需要食物和植物需要阳光的原因。对于遗传算法，能量成本同样存在。计算机需要用电。从内存和硬盘读写数据都需要能量。另一个永远的需要是维持。总的来说，部件越多，组织程度越高，维持结构所需的成本就越高。用于维持的能量将低熵维系住。

熵度量无序，有组织的系统有变得无序的自然趋势。热力学的中心原理是能量可以用来抗衡熵增。保持桌面整洁就是典型的例子。我们需要耗费能量才能让桌子保持干净；否则纸张、书本、笔、文件夹和备忘贴就会堆起来。不努力维持，任何有组织的系统最终都会衰退于无序。本质上，复杂引擎的循环之所以能对抗熵增就是因为它能从无序中计算

提取目的性有序。有序的增加表现为目的性结构的形式。随着进化系统累积越来越多的结构，它们也越来越容易被随机变化损坏，需要越来越多的能量维持。任何变得越来越复杂的系统都面临一个现实，就是维持所需的能量成本最终必然会等于继续增加复杂性所能带来的收益。

时间成本不那么明显，但非常重要。如果从响应挑战的角度来看实体与环境的关系，那么在许多情形中优势都来自于更迅速的响应。当狮子遇到羚羊，速度就是一切，而且不仅是奔跑的速度，还包括决策的速度。时间成本更微妙的一面来自于复杂的响应比简单的响应更能应对某些挑战，而复杂的响应需要花时间计算。当其他条件一样，就需要在响应的复杂性和响应的速度之间进行权衡。羚羊也许可以依靠聪明逃离狮子，但如果思考的时间太长，结果就不太可能对羚羊有利。

计算复杂行动和复杂结构的时间有一个容易忽略的方面是需要的时间有时候可以提前支出。我们可以对比一下狮子与羚羊遭遇和狮子与有准备的人遭遇。人类跑不过狮子，但可以带武器。枪或长矛需要很多时间制造，但这个时间是花在遇到狮子之前而不是之后。武器内在的复杂性以可能有用的结构的形式制造和保存下来以备不时之需。当这样的情形经常发生，就会去寻找需要时间来创建但能够在需要时快速响应的结构。复杂性允许快速响应但需要深入计算的这方面就是深度。

空间也是很重要的资源，但在物理世界通常不受限制。在计算机科学中空间表现为有限的内存，但前面曾说过，受限的空间往往能通过花费更多时间进行补偿。

深度有何意义？

深度度量的是计算对象内部关系的复杂性 [15] 。基本思想是这样：内部关系越复杂，从没有这些关系的输入计算出有这些关系的输出所需的步骤就越多。这个特性被称为“慢增长”，意思是要快速创造深度对象只有用深度输入才能做到。用简单（浅）对象创造复杂（深）对象需要大量计算。这条规则没有捷径。由于物理过程也可以被视为计算，因此物理对象也必然具有类似深度的属性。

这个概念可以直接扩展到智能体之间的互动。当挑战和响应被视为

计算，深度挑战和深度响应都需要用更多的计算才能创造出来。

慢增长特性有力地解释了为什么深度实体可能更受一些进化情境青睐。在自然界中生物面临着数不胜数的取决于时间的挑战。由于可以预先准备，并将结果以深度结构的形式存储，因此就有可能预先创造或计算出结构，在需要时就可以产生快速的深度响应。在野外携带枪支就是一个例子。根据定义，浅响应必定要么是随机的，要么是高度重复的，而随机或重复的行为通常不是最有效的。因此当快速的深度响应能带来优势而附带的成本又不是太高时，选择就会青睐于能产生快速复杂响应的深度结构。动物大脑和复杂的代谢网络很有可能就是深度生物结构的例子，由青睐能有效应对深度挑战的快速响应的长期选择创造出来。这些结构同时也能增加控制能力。

如果简单响应就能够有效应对挑战，深度计算就是对时间和能量的浪费。可预测的挑战通常只需记忆响应（简单计算）。完全不可预测的挑战则根本无法计划应对。这时最佳选择就是随机响应（也是简单计算）。介于两者之间的大量挑战的最佳应对是深度计算行为。生物界有大量这种例子。以猫抓老鼠为例。猫需要食物，否则就会饿死。老鼠躲在洞里，主要在夜间活动，这样很难被发现。为了响应这个挑战，猫进化出了能快速响应的肌肉，善于夜视的眼睛，优秀的听力，以及能够预测在哪里能找到老鼠的大脑。

虽然有一些挑战只需要简单响应，但许多挑战都需要深度计算才能更好地应对。尤其是深度挑战的最佳应对往往是深度响应。但深度挑战从何而来呢？在生物世界，生物面对的物理环境一般不会很复杂。物理环境的大部分方面要么随机要么重复。例如白天黑夜的循环，冬冷夏热；与气候相比每一天的天气相对随机。如果生物只需要应对物理环境，我们可以预计它们不会进化出特别复杂的结构。

而如果环境中有相互竞争的进化智能体，则是另一回事。每个实体都面临着其他实体的挑战。每次行动也都会形成对其他实体的挑战。以猫和老鼠为例。在长期的竞争过程中，老鼠越来越擅长躲避猫，猫也越来越擅长捕捉老鼠。在进化的智能体挑战和响应的过程中，选择总是青睐于更有效的响应，惩罚那些落伍的。随着道高一尺魔高一丈，群体整体上的“成功”相对来说也许基本没变。这就是所谓的“红皇后”原理，由芝加哥大学的利·范瓦伦首先提出[16]。高效的重要一面在于及时性，当

更高效的响应既需深度计算也对时间敏感时，挑战和响应就会趋向于越来越深。这个正反馈环决定了深度结构的进化创造。如果我们接受深度作为复杂性的一方面，协同进化就必然会导致其增长。

在非生命进化系统中，如果进化实体之间存在竞争，并且响应时间很重要，则必然也存在同样的反馈环。因此可以预计，如果用深度作为复杂性度量，当进化系统的进化实体之间存在竞争时，除非响应时间不重要或者与复杂性相关的能量成本太高，否则复杂性一定会增加。

有证据吗？

在科学中，一个预测要有价值必须是可以验证的。对于深度是协同进化系统产生的复杂性的重要方面这个假说也同样如此。深度的计算概念的一个大问题是它是不可计算的——它是无法度量的度量！第2章介绍的算法信息是另一个具有同样让人困扰特性的计算度量。幸运的是，两者都可以被近似。

从概念上，深度刻画的是从简单输入计算某物的难度。相反的计算过程是压缩，从长输入得到简短的文件，而且不损失信息。压缩是很常用的软件，广泛应用于文件存储和网络传输。当文件被压缩，通过解压可以将文件还原出来。JPEG和ZIP都是压缩文件。压缩文件在使用前必须解压。“压缩深度”是对压缩一个数字对象的难度的可计算度量。大部分学者认为，很难创造的对象（例如深度对象）也很难被压缩。

1997年，当时还在艾奥瓦州立大学读计算机科学研究生的詹姆斯·莱思罗普设计了几个计算机实验，让二进制实体在由进化实体组成的环境中进化。他的研究与“有限状态自动机”有关，自动机是可以被视为智能体的计算机实体，因为它们能根据当前情形和过去的行为记忆作决策。莱思罗普让这些智能体进行反复的囚徒困境博弈，然后用遗传算法根据成绩让它们进化。不出意料有限状态自动机智能体进化出了越来越好的策略。由于这些智能体同时也是计算机文件，因此莱思罗普可以度量它们的压缩难度。他采用的具体度量方法是基于被广泛使用的L-Z压缩法。图12.5展现了实验过程中智能体的压缩深度的增长。可以看到，在3500代中智能体的压缩深度稳步增长，没有明显的变缓。

对这个实验很直接的解释是L—Z压缩深度度量了智能体的组织复

杂性的多个方面。牵涉的组织层次越多，重新编码成更紧致形式的计算需求就越大。计算深度度量的是完成指定的压缩所需的计算资源（在这里是内存）的量。这个实验的一个重要特点是智能体（有限状态自动机）没有变得越来越大（需要更多计算机编码），而是似乎组织得越来越巧妙。这个实验可重复，而且用不同的数据结构和不同的选择任务也观察到了同样的现象[17]。观察到的变化可能与前面讨论的整体适应有关。无论怎样，这些发现都说明了在协同进化的情境中计算实体内部结构的变化无法用香农或算法信息的概念解释。其他协同进化系统也发现了类似的现象。

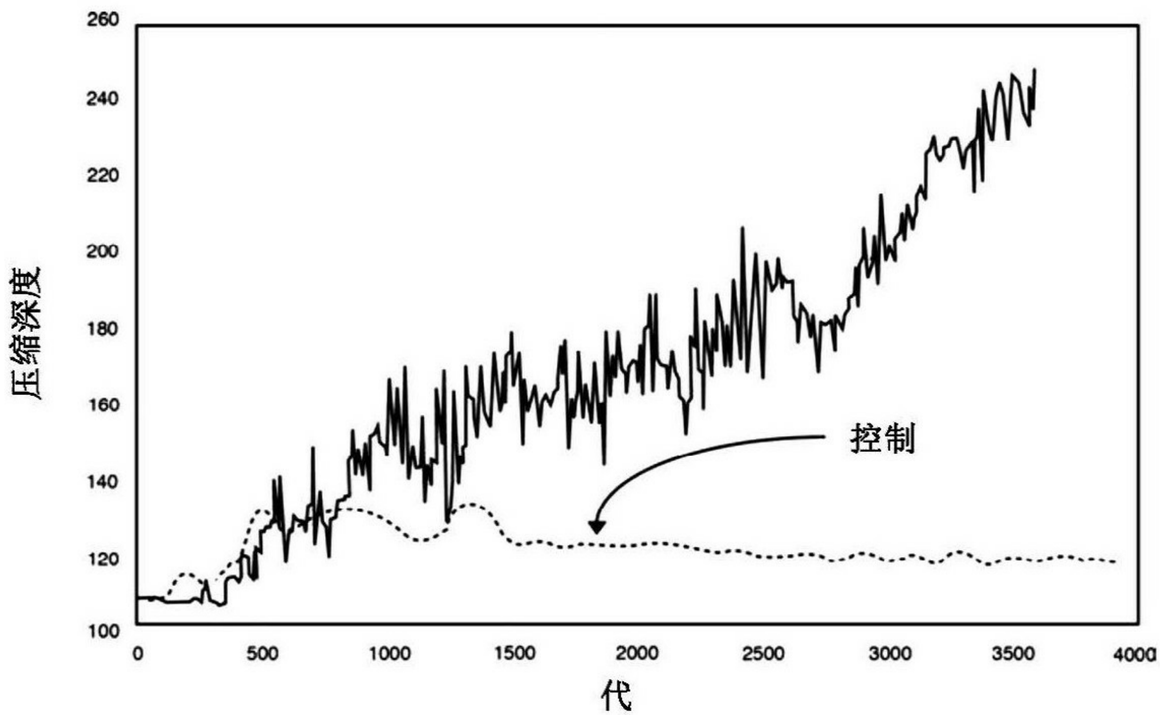


图12.5　适应过程中L—Z压缩深度的增长。经莱思罗普许可复制（参见注释17）

如果可以比较远古生物与现代生物的DNA序列的压缩深度将会很有趣。可惜化石中的DNA无法留存下来。而可以进行实验的（例如冰河时期的人类或猛犸的DNA）时间又太短，无法检测出差别。终极实验是比较生命史初期生物的DNA与很久以后的那些生物的DNA的深度。当然这是不可能的，因此我们只能通过分析进化算法的结果来推断生物世界的情况。

深度是计算机科学的度量，还没有被研究进化的非计算机学家广泛

认识，当然部分是因为它无法度量。莱思罗普和阿什洛克的实验清楚表明了在一些协同进化系统中存在的复杂性增长无法用传统的度量进行刻画。研究深度的学者对这个特征有各种说法：“对象内在的值”，“计算有用性”，“对象明显的计算历史的量”，“冗余的组织”，“精巧性”，“组织中存储的计算工作量”，以及“通过计算嵌入对象中的组织”。这些对抽象数学概念的各种描述的共同线索是信息可以被计算过程组织以创造输出结构，当它们被作为进一步计算的输入时会表现出性能的提升[18]。展示性能的一个重要方式就是速度。在许多背景下快速复杂的响应都很有用。深度结构使其成为可能。

这个度量的存在带来了一种可能，当人们感受到复杂性时，他们部分感受到的是深度。深度对象的一个特征是它们必然有很长的历史。一棵很老的树，人的大脑，以及成熟的生态系统都是需要很长历史的例子。它们都无法被快速创造。这就是深度的标志。也许正是对深度的感受使得我们喜欢欣赏有很长历史的事物。

第13章　过去和现在

复杂引擎如何启动？

刚开始学习进化的学生常问的两个问题是：生命是如何开始的？以及未来会怎样？这些都是很难的问题。复杂引擎的思想能为这些问题带来新的视角。我们从起源开始。

回顾一下，我们探讨了4个现在在地球上运作的复杂引擎的例子：生命、哺乳动物的适应性免疫系统、社会演变（宗教、科学和技术以及经济作为子例）以及进化算法；人类学习是可能的第5个例子。引擎的所有这些实现都需要能编码信息和执行计算的物理系统。我们的下一个任务就是比较各个系统的起源以找出共同特征。

生命的起源

关于生命的起源有很多理论，但都无法从科学上让人完全满意，主要是因为直接证据太少了。我们对一些事情很有信心。首先，生命在地球上的出现从原则上只有3种可能途径。生命可以是：（1）来自宇宙其他地方；（2）通过物理和化学规律以外的某种力量的作用在地球上出现；（3）通过自然的化学和物理过程在地球上出现。第3种可能又分为有合理可能性和基本不可能（例如奇迹）的情形。目前没有证据表明生命起源于其他地方，但如果在火星或木卫二上发现了基于DNA的生命将会彻底改变这一点，而违背已知物理和化学定律的起源不在考虑的范围。因此，科学家们只有一种选择，就是研究地球上有可能存在的化学场景，当然NASA也一直在资助在火星和太阳系其他地方探测生命的实验。

化石和地质证据表明生命在地球上出现于35亿年前，但不会早于42亿年前，有微弱证据表明可能是38亿年或之前。留存下来最古老的岩石大约是在38亿年前形成，虽然名为锆石的晶体更古老。这个时期地球上的环境相当不稳定。标准的太阳系模型表明当时的太阳亮度只有现在的

75%，这意味着海洋会结冰，但二氧化碳和甲烷等温室气体的存在以及温度更高的行星内核可能会让地球表面足够温暖，让海洋保持液态。有许多火山释放气体。事实上，大部分早期地球模型都将火山喷发作为大气的大部分或全部来源。氢、碳、氧、氮和硫都有，但对于这些元素的化学形式还有争议。

最大的未知是大气的氧化态。现在的大气中有很多氧气，但科学家们都认为这是光合作用的结果。生命通过分解水（H_2O）和释放氧气（O_2）维持大气中的氧分子，所需的能量来自阳光。氧分子具有化学活性，如果没有光合作用很快会从大气中消失。在氧化条件下，碳会成为二氧化碳（CO_2），氢会成为水（H_2O），氮以氮气（N_2）或硝酸根离子（NO_3^-）的形式存在，硫则成为硫酸根（SO_4^-）。硫和氮溶于水或束缚在岩石中。与氧化相对的是还原。原子与氢而不是氧结合。这种情形下，碳会成为甲烷（CH_4），氮会成为氨（NH3），硫会成为硫化氢（H_2S），氢会以分子形式（H_2）存在，氧则存在于水（H_2O）中。介于两种情况之间的是水、氮分子、二氧化碳或一氧化碳（CO）以及硫元素（固体）。其他化合物包括甲醛（CH_2O）和氰化物（HCN）都有可能。

通过深入研究地质过程和仔细分析最古老岩石的化学，地球化学对早期大气成分的认识被不断改变。目前的模型倾向于认为是有少量氧分子的中性或适度的还原性条件（H2 O、N2、CO2，可能有CO、S和/或H2 S）。在火山口附近，局部可能具有很强的还原性条件，现在依然是这样。通过太阳光、干湿循环、频繁闪电以及大量火山活动，会发生大量由氢、碳、氮、氧和硫参与的化学反应，形成大量碳基分子。土卫六上现在可能正在发生类似的过程，土卫六的大气含有甲烷和氨，具有很强的还原性条件。由于存在基于碳、氢和氮的复杂有机分子，它的大气呈现出朦胧的橙色。土卫六很冷，因此氧大部分封存于冰冻的水中。

被引用最多的生命起源理论可以追溯到达尔文1871年写给约瑟夫·胡克的信：

经常有人说，如果第一次产生生命的所有条件都存在，它就会一直存在。但如果（噢！多么大的如果啊！）我们设想在温暖的小池塘里，有各种氨和磷酸盐、光、热、电，等等，蛋白质通过化学方式得以合

成，并产生更复杂的变化，在今天这些物质会马上被蒸发或吸收，使得生命来不及形成。

今天，达尔文说的“温暖的小池塘”更常见的说法是“原始汤”，由氨基酸、糖、核苷酸和其他各种有机物溶解在水里进行反应形成的浓汤。这个构想还有许多问题没有解决。首先，虽然大部分关键化合物在实验室中在合理的条件下可以产生出来，但所需的条件对于所有化合物并不是一样的。因此只能设想不同的化合物是在不同的池塘中产生出来，然后碰巧以正确的顺序组合到一起。另一个问题是，要产生出聚合物（蛋白质、多糖和核酸都是聚合物），“汤”的浓度要足够高，这在海洋里是不可能的，因为当水太多时，作为生命特征的有机物的化学性质不稳定，很快就会分解。

为了解决这些问题提出了许多构想，包括在干涸的池塘里浓缩，在冰上形成的小水窖里浓缩，催化黏土表面的特殊化合，与热岩接触的特殊环境中，比如在水底火山附近，在沉积物（泥巴）中化合，以及在太阳星云中化合，然后被彗星和陨石带到地球上。所有这些想法都开启了新的探索维度，但现实是目前还没有人能证明在实验室条件下如何生成所有所需的化合物，并且合理地解释这种条件在早期地球曾经很普遍[1]。

不考虑化学细节，深入研究过这个问题的人都认同生命的形成有两个关键。需要出现能自维持的连锁化学反应网络，也就是代谢的前身，以及各种化学反应的催化和调控，组成受指令控制的网络。需要催化是因为许多现代代谢反应在常温下都不会发生。在所有现代生命形式中，催化和控制都是通过蛋白质酶完成的，而蛋白质酶的结构是由DNA编码的信息决定的。复杂引擎能对现代生命起作用是因为这种信息可以复制，并受到累积性选择。

专家们对于代谢和指令谁先谁后一直有很大争议。先有鸡还是先有蛋？一些人认为信息编码分子的自复制肯定是种子事件。因为一旦实现了自复制，对指令进行扩展，实现对其他反应的催化（包括自复制反应）就不是很难的事情了。多年来，对DNA编码蛋白质和蛋白质参与合成DNA的认识使得这个难题被认为是问题的关键。让人吃惊的是，这个问题在20世纪80年代就从原理上被解决了，却并没有解决生命起源问题。重要的发现是，与DNA非常类似的信息携带分子RNA能催化化学

反应。因此在原理上可能存在RNA分子能催化自身的合成。还没有化学家能生成出这样的分子，但如果这个分子能够通过随机化学反应合成，并位于核苷三磷酸（RNA的化学前身）的浓汤中，它就能自发复制自己。由于这种假想的分子能在其结构中编码生成更多RNA分子所需的信息，它也就能自动进入复杂引擎的循环，它的后代必然会很快变成越来越好的复制者。这个思想使得在20世纪80年代出现了一种生命起源早期的构想，称为“RNA世界”——自复制RNA分子的池塘，RNA分子相互竞争空间和核苷三磷酸。在选择压力下，这些RNA分子（通过进化过程）逐渐“学会了”编码蛋白质酶，能催化更简单的分子形成核苷三磷酸，并最终产生出现代细胞生物化学网络所需的所有生物化合物。在这个构想中，复杂生物化学网络所展现的结构是复杂引擎作用于化学系统的自然结果。

这是一个很有吸引力的构想，但虽然最近有一些突破，还是没有人能令人信服地解释所需的核苷三磷酸是如何产生出来并形成足够高的浓度从而合成出RNA链，也没有找到能自复制的小RNA [2] 。另一方面，自复制分子作为关键的第一步的总体想法仍然很有吸引力，一些科学家正在不断努力寻找能使其发生的条件。

代谢在前的构想有更久远的历史。最近的一个版本是斯图尔特·考夫曼提出的 [3] 。其中的思想是，关键的第一步不是自催化的复制分子，而是自催化自维持的化学反应网络。他称之为“自催化集”。这似乎很有可行性。在一个多少有些随机的化学反应系统（汤）中，会（随机）形成一些简单的聚合物。单体能形成二聚体（两个单体连到一起）和三聚体（三个单体连到一起），二聚体和三聚体又形成四聚体、五聚体、六聚体，等等。随着连接的单体数量越来越多，可能产物的数量迅速增加。最初，这些聚合物的浓度很低。

一些聚合物有一个特点是能催化特定的化学反应，例如一些蛋白质和RNA就有催化作用。因此不难想到在反应汤中会有一些小聚合物，而其中一些能催化化学反应。至于是催化何种反应并不重要。催化反应的化学产物会迅速增加（因为催化），从而形成新的反应和新的聚合物。一旦形成新的催化，化学网络的许多细节就会改变。只要有一个催化反应能间接导致其催化物浓度的增长，在反应产物和催化物之间就能建立正反馈环。最让人感兴趣的是小聚合物通过加速比如说特定的二聚体与三聚体结合，从而催化大聚合物的形成。小聚合物催化这种反应的可能

性已经被证实。例如，很简单的聚合物双甘氨肽（甘氨酸是最简单的氨基酸，双甘氨肽由两个甘氨酸结合形成）在高浓度盐水中能催化一些氨基酸和小缩氨酸的聚合，这在干涸的池塘中有可能实现[4]。产生的新聚合物又能与其他小聚合物结合产生出更大的聚合物，每种新产物都具有催化某种反应的可能性。如果小聚合物催化形成大聚合物的概率足够高，这个过程必然导致越来越多的正反馈环形成。

整个想法看似有点不现实，但考夫曼和他的同事已经用计算机模拟证明了它确实可行。通过假设随机选取的聚合物催化聚合物反应的概率，当最初的反应汤中化合物种类足够多之后，就会出现“催化闭环”，从而使得网络只要有简单的化合物原料供应就能自我维持[5]。在第9章就是用自催化自维持的生物化学网络来定义生命。如果考夫曼的观点是正确的，今天细胞中的生物化学/蛋白质组/调控网络就是从这些早在第一个细胞甚至RNA或DNA出现之前就形成了的自催化生物化学网络不断演化而来。在很早的某个时候，这个网络逐渐加入了指令（很有可能是以RNA的形式），指令可以更有效地管理和累积目的性信息。指令也更容易进化。

这个构想有一个吸引人的特点是普遍性。所需的不是特定的化合物的功能，而是催化闭环，只要满足一定的普遍性条件以及化合物种类数量足够多就可以。不过同期它假说一样，要让人信服，必须能在实验室展示具体的化学过程。在已知的早期地球条件下，到底哪种催化聚合物是可行的？多肽（短蛋白质）是一个好的候选，因为氨基酸可以在各种模拟的早期地球条件下产生，短多肽具有催化活性，并且多肽是现在生命所使用的催化剂（细胞酶是长多肽）。但还没有人能在实验室实现基于人工多肽的自催化集，在实现之前，有理由持怀疑态度[6]。

自催化模型的一个有趣特点是，聚合物就其本质来说，在序列中编码了所催化的反应的信息。因此，聚合物自催化集在其多个聚合物中分布式编码了其自身繁殖所需的信息。选择将有力地塑造这样的系统中涌现出的动力学和细节。如果某些网络反应能合成出核苷酸和核苷酸多肽，那么从自催化集到RNA的清晰路径就不难想象了。

RNA和自催化集在传统上被视为生命起源的两个不同选项，经常被认为一个是“信息在先”，一个是“生物化学在先”。但就如我们看到的，它们并不一定冲突。如果催化剂引入了RNA，RNA的世界就能从之前的

自催化化学网络中产生出来，并且这个网络还不断给它提供核苷三磷酸。

总的来说，信息在先模型需要信息分子的随机形成，而且这种信息分子要刚好能催化其自身的复制。只要是能自复制的分子就会自动进入复杂引擎循环（假设其复制还可以再完善）。由于RNA序列能催化特定的化学反应，因此不难想象，这种自催化反应如果存在的话，通过进化将可以构建出生物化学网络，在最初的原料耗尽后能继续提供复制所需的原料。

与之对比，生物化学在先模型则猜想，生物化学网络在将分布式催化和调控信息升级为专门的信息编码分子（RNA或DNA）之前，已经变得相当复杂。根据这种猜想，有了专门存储和利用信息的分子的加入，复杂引擎才会全力开动，但在此之前选择已经作用于已存的生物化学网络。根据这种观点，最初的信息分子不是自复制，而是同自繁殖网络的其他化学成分一起合成。

逻辑分析无法判断哪一个思想更接近真相。只有在实验室中建立起合理的早期地球地质化学条件，通过实验证实，才能最终判断它们（或其他理论）是不是正确。

生物化学在先的思想有一个吸引人的特点是，在有明确编码指令信息的分子出现之前，选择就起到了重要作用，并且选择绝对不要什么复杂技术。无论是生命还是非生命，选择在自然界都广泛存在。生物化学模型也不怎么依赖特定的低概率事件。一旦具有特定属性的系统足够复杂，它就会自组织成自催化集。

对于信息在先模型，通过简单的概率计算就能看出最初的自复制分子所能编码的目的性信息的量多么有限。估算是这样：假设系统随机生成携带信息的分子的速率是每秒一个。这是活细胞中大肠杆菌RNA聚合酶转录一个基因所需的时间。原始的复制系统基本不可能有这么高的效率，因此这是保守估计。然后假设每秒都有大量分子合成。为了论证需要假设一个具体的数字，假设为6×10^{23} ，也就是化学中的1摩尔 [7] 。这个量很大，会产生很多的单体分子。让时间持续1亿年（2×10^{15} 秒）。这么久大约会产生10^{39} 个分子。

如果这些分子都是RNA（有4种化学基作为信息编码字母），则可以编码2^{130} 比特信息。也就是说，这个系统1亿年可以穷尽搜索130比特宽的信息空间。这意味着经过1亿年大部分65碱基长或更短的RNA会被生成出来（每个碱基2比特），但大部分更长的分子不会出现（随机合成）。如果我们的产物用完地球上所有的碳原子（很荒谬的假设），能探索的信息空间也不会超过200比特宽。由于长度和可能序列的数量呈指数关系，因此出现在早期地球上的具有特定意义的聚合物序列的长度不会超过100个单体分子，50是更合理的上限。因此，通过简单的数学计算就能得出结论，最初的信息分子如果是随机生成的话，不可能很长。

100比特（50个RNA单体）的信息还不及一行字。“煮蛋3分钟然后加盐和胡椒”的指令就有大约100比特信息。信息在先模型必须要能够构造出50单位的聚合物，并折叠成能催化自身复制过程的形状（提供合适的简单前身分子）。目前还没有发现这样的分子；但也不能下结论说不存在。而对于RNA分子连接酶的活性则已经了解得很清楚。连接酶能将两个聚合物连接成更大的聚合物。RNA连接酶可能在最初的自复制RNA的产生过程中起到了重要作用。生物化学在先模型的一个主要优点是最初的编码信息不需要具有自复制功能，只要能改进网络性能就可以。在生物化学在先模型中，概率问题似乎是被逐步克服，不需要依赖高度不可能的步骤（最初的催化闭环事件的概率目前还无法计算，专家目前的估计范围从“基本肯定”到“基本不可能”都有！）。总结一下，生命起源的关键事件是信息编码分子或信息编码自催化网络的自复制，目前还无法判断哪个更可信。两者都是基于选择。

适应性免疫系统

再来探讨一下适应性免疫系统的起源。所有生物都有防御微生物攻击的分子机制。没有它们，任何动植物或菌类都会很快成为细菌的食物。防御机制有很多，其中大部分被归于“先天免疫”。防御机制包括识别和摧毁外来DNA和RNA，识别细菌表面特有的多糖和蛋白质，分泌抗菌肽，以及在动物体内巡逻，吞噬可疑物质的阿米巴样细胞。动植物的存活就是这些机制成功的证明。先天防御机制有一个固有缺陷，微生物进化出反制措施的速度要比缓慢繁殖的宿主进化出新的防御措施的速度快得多。大约4.5亿年前，脊椎动物出现后不久，一种新的策略出现

了，脊椎动物能在体内进化出新的防御系统，速度同微生物的进化速度一样快。我们称之为适应性免疫系统。

这个系统的进化是渐进的，包含许多中间步骤。其运作的基本原理在第7章已经介绍过。针对入侵者专门定制基因以生成新的抗体和新的T细胞受体是适应性免疫与先天免疫的主要区别。这个系统依赖于两个不同策略的组合使用，重排现有的抗体基因片段形成新的抗体基因，以及只选择那些生成的蛋白质能与病原体结合的新抗体基因进行复制。DNA序列分析表明，执行基因重排的酶与原始脊索动物身上发现的转座酶有关联，脊索动物是与脊椎动物最接近的生物[8]。转座酶剪切DNA短链并将其插入DNA上新的位置。被转移的DNA称为转座子。如果转座子包括编码负责转移它的转座酶的基因序列，这个部分（转座子和生成的转座酶）就可以视为能通过将自己插入宿主DNA上新的位置来进行复制的DNA寄生虫。数据表明，有突变使得转座酶可以与编码抗体片段的DNA结合，而不是其自身的DNA相结合。这个错误被不断强化，从而可以产生出许多不同的抗体序列，而不用在DNA上附上数以百万计的抗体基因。这个机制要起作用，还需要有细胞能够生成抗体与外来抗原结合，从而起到选择和放大的作用。有可能在转座酶基因突变之前，原始脊索动物已经具有了一些抗体基因，以及判断哪一个与入侵者最匹配的识别和放大机制。

细胞分裂和选择的一般机制从多细胞生命出现起就存在了，比免疫系统要古老得多。细胞的选择取决于细胞表面与其他蛋白质（或多糖）结合的蛋白质。如果结合错误而不是正确的蛋白质，细胞就会被淘汰。免疫系统的选择复制与此类似，不同之处在于，选择以及随后的快速复制的刺激，是由与环境蛋白质（或多糖）的结合触发，这些不属于身体自有的常规大分子库。从受体内正常分子刺激变为受外来分子刺激，这个简单的变化可能是通往适应性免疫的第一步。

如果这个设想是正确的，选择就先于免疫系统中复杂引擎的全力发动。与生命的起源相比，适应性免疫的产生相对简单一些，这是因为已经具备了自复制信息编码系统，并且适应性免疫系统所需的关键要素也已经因其他目的存在。现代免疫系统的关键要素包括细胞分裂（以及DNA编码蛋白质）、抗体、T细胞受体、让结合蛋白质快速产生变异的重组系统，以及根据细胞表面蛋白质与特定分子的结合进行选择。在免疫系统出现之前细胞和蛋白质就已经存在，而且就如我们看到的，多细

胞生命都根据细胞膜蛋白与其周围物质的结合对细胞的谱系进行选择。

大脑的发育

第7章曾简要介绍过，多细胞生物的构建是通过对分裂的细胞进行特定的选择。通过细胞的此消彼长发育出新的结构。在大脑中选择策略还作用于亚细胞层面，决定神经元之间的连接细节。神经元之间会建立随机连接，然后一些连接被选择留下，另一些则退化，从而建立起大脑的连接细节。神经往外生长连接远处的目标，通过细胞表面传递特殊的分子信号。分子信号促使神经元继续往前生长，形成突触，或者退缩然后往其他方向探索。探索和选择结合到一起，抽取出随机性所固有的创造性，而不用复杂引擎的全力推进。在大脑发育的过程中，没有对描述连接细节的指令的复制和修改，并不存在信息体来指定所形成的模式的细节。发育系统的确会在结构中记录过去的选择，但在大脑发育过程中，这种信息并没有被反复复制和修改以达到不断改善的目的。大脑连接展现了受巧妙规则引导的反复选择的力量。当然，这些规则是编码在生物的DNA中，通过漫长的生物进化创造出来。

在远古时期，动物生命的策略发生了改变，从预先确定身体和大脑改为通过随机的尝试和非随机的选择来塑造结构。科学家们喜欢用线虫来研究身体发育的细胞和分子机制。这种1毫米长的小蠕虫通常生活在温暖的土壤中，但在细菌培养皿中有喜欢的食物也能适应。雌雄同体的成虫有959个细胞，在身体上的位置基本固定（除非发生了变异）。发育是确定性的：如果移除一个胚胎细胞，成虫就会缺失由这个细胞发育而成的一部分身体。神经系统也是一样。成虫有302个神经元，出现的位置具有很高的可预见性，相互的连接也高度可预见。

线虫的DNA编码了大约20000个基因，与人类的DNA相差不大（人类大约22000个基因）。显然，人类要复杂得多，然而却只需增加很少的基因就能产生出人类。这说明了什么？我认为答案在于我们身体的形成很好地利用了选择的力量。在人体中，没有哪个细胞的位置是精确预先确定的。如果缺失了某个胚胎细胞，其他细胞的后代会接替其功能。人体在单个细胞层面上很不一样，在组织和器官层面上却或多或少是一样的。这种发育策略与线虫不一样。显然如果细胞的位置和功能都明确指定，则需要20000个基因才能构造有1000个细胞的生物。人体有

10000000000000（10^{13}）个细胞。如果所有细胞的精确位置都由DNA中的指令预先确定，要实现这样的复杂性是根本不可能的。在远古时期从确定性策略到基于变化/选择的系统的转变使得我们成为可能。这也使得复杂的大脑成为可能。

思维

对思维的起源知之甚少，但线虫应该不太可能有思维。它的302个神经元通过决策提高获得食物、交配和避开不利环境的机会，但要说这种简单活动是思维很难让人信服。据我们所知，线虫的决策是固定连线的，有点类似于计算机中运行的人工神经网络（第6章）。鸟类和哺乳动物等高等动物显然具有思维。只要观察过动物如何解决问题，比如松鼠或乌鸦如何从“防鼠”鸟食器中获取食物，就会相信这一点。狗的行为也极具智慧。当然，我们无法知道动物到底在想什么，因为我们无法与之交谈。我们需要交谈才能了解他人的思维。不过我认为如果动物的表现和行为很像是在思维，就没有理由拒绝接受将动物的行为解释为它们的确是在思维。它们可能不像人类思维那样有精细的场景构建能力，也肯定不会用人类语言思考，但没有科学理由质疑鸟类和哺乳动物具有思维；显然它们也学习。甚至昆虫和较高级的软体动物也会学习。就连只有302个神经元的线虫也能学会趋近或避开与食物有关联的味道、气味或温度。

在第10章我曾推测思维之所以可能是因为大脑具有由复杂引擎塑造的计算策略。如果将学习解释为复杂引擎在大脑中以微秒级的速度运作，则需要很久以前在动物的进化史中偶然出现了一种学习策略，原始大脑对感知输入同时产生多种解释，然后选择与其他感知和记忆数据最匹配的解释。一旦实现这一点，只需稍加改进就能引入多回合的修改和选择；即复杂引擎。这是改进不完美的内部世界模型很自然而直接的方式。

无论线虫的大脑中编码了怎样的原始世界模型，它都必然源自标准而缓慢的生物进化试错过程。如果你很小，又生活在土壤中，你的世界在大部分时间里就具有很大的可预见性，但如果你很大，生活在海洋或陆地上，你就会不断面临新的环境。实时适应（而不是进化适应）必然会带来极大的优势。通往思维的进化的第一步可能出现在动物进化相当

早期的时候。如果对于周围环境有一个心智模型将非常有用。没有这个模型就只能用基本随机的行为应对环境。好的模型能让人知道到哪里躲藏和寻找食物。即使是很初级的思维，心智模型也是必需的。问一些将自己置于不同位置或面临行动选择的环境的问题。我应当到A还是B地点寻找食物？我应当躲在X岩缝还是Y岩缝？一旦具有了这种能力，大脑所需的就是在不同的心智模型之间进行选择。

思维的进化与针对特定疾病的抗体的产生有一个重要的区别。抗体和思维都是基于对之前编码的信息的复制和修改。身体结构基本会不变，大脑中的连接模式则是暂时性的，可以有多种形式。大脑的生理结构具有很高的冗余性，很适合产生相似却又稍有不同的连接模式，根据与其他活跃神经模式的匹配程度进行选择。

学习需要有选项，而选项需要有选择的机制。当选择固定下来，就发生了学习。由于身体反应（运动、记忆、思维）是由神经活动模式决定，而神经活动模式又编码了信息，因此即使是简单动物的学习也涉及对信息编码模式的选择。不难想象大脑冗余会以这样的方式组织，最初作出和记住的决定可以通过反复的修改和选择的内部机制得到改善。因此，思维的根源与学习的根源可能密不可分。动物神经系统最简单的是珊瑚虫纲：水螅、珊瑚和水母。没有发现这种初等动物有联想学习。这些动物也没有中枢神经系统，虽然它们有神经细胞，能对刺激做出正面或负面反应；但它们似乎没有对间接线索进行联想的能力。所有有脑的动物似乎都能进行初级的学习。

对于让原始大脑能够学习的生理结构的革新还知之甚少，但无论是什么，学习必定是由基于细胞通信的时空模式的信息编码系统所设定。这个编码策略完全不同于基因编码。随后的关键步骤可能是能够提供选择的中枢神经系统的形成。这可能发生在寒武纪物种大爆发的早期。

社会进化

如果能够用大脑编码和修改信息，又能大致相互交流，还能选择性地记忆，则社会的进化就是不可避免的。当思想和概念通过人际传递，也就是在复制信息。由于我们会选择记住或忽略接收到的信息，有时候还会对听到和读到的添油加醋，可想而知社会编码的总体信息会不断进

化。真正的问题在于为什么原始社会的进化如此缓慢。

社会进化的起源与交流的起源密不可分。社会进化不仅限于人类：打开车门的技巧很快在约塞米蒂国家公园的熊中普及，却没有传到大部分野熊种群。年青的熊显然从年老的熊那里学习技巧。许多黑猩猩和猩猩群体使用不同的技巧获取食物。只有进化出了一种新的信息交流形式——语言——之后，科学和宗教等精致的社会制度才成为可能。

今天的社会进化背后的信息不仅限于说的话，还包括书写文字、图表、电影、音乐和电子设备解读的二进制码；但要对社会起作用，信息就必须有人类的大脑处理的阶段，编码为神经连接模式。社会进化是基于人类学习的复杂引擎的产物，而人类学习又很有可能是基于大脑的复杂引擎的产物。语言的进化是因为它能为掌握了语言的人带来好处，比如准确描述机会或危险。但一旦语言产生了，人们就会开始分享创新、喜好和信仰。当然个人会选择接受一些又忽略另一些。对于那些最善于交流的群体来说，这会带来极大的好处。人类社会进化的起源与一种新的信息编码方式——语言——的起源密不可分。对于语言的起源知之甚少，但很显然需要新的大脑结构（DNA进化的结果）。

进化算法

进化算法是运行在计算机中的复杂引擎的实例，它们是人类的技术创新，本身也是社会进化的产物。进化算法的起源有很详细的文献记录。1940年，最早的一些程序员就开始用程序模拟生物进化的某些方面。结果成功了。DNA不是必需的，细胞也不是，生命的所有具体细节都不是必需的。要在计算机中人为构造出能进化的系统，所需的只是编码的信息，复制信息的机制，在复制过程中引入适度的变化，以及根据一致的标准选择部分信息进行复制和修改的机制。显然，这些算法的来源是人类的天赋和对生物的模仿；另外当然，算法少不了计算机。由于进化算法是预先设计的，它们的起源没有什么神秘。人类发明了它们。

在更深的层面上，复杂引擎在计算机上的实现要有编程语言才有可能，这是编码和使用信息的新途径。因此，人类社会的进化需要人类语言，而计算机进化算法则需要编程语言。计算机是社会进化的产物，社会进化又是能够交流、思维和学习的神经系统的产物。思维的大脑本身

又是基于DNA序列所决定的复杂发育系统的背景下实现的随机尝试和选择才成为可能。而DNA序列本身又是在整个生物层面上无数次实验和选择的结果。因此这是复杂引擎的嵌套实现，每一个都是基于一个更古老的进化系统。唯一不是（明显）来自更古老进化系统的就是生命本身。

我们所讨论的所有的复杂引擎的例子，除了适应性免疫，都有不同的编码信息的方式，信息被用于生成某种东西或使得某种事件发生。表13.1对此进行了简要总结。

其中除了最后一个，复杂引擎在启动之前都有一个系统利用了选择但没有编码和复制信息的机制，而且在每个例子中，除了适应性免疫，都是新的编码和利用信息的方式使得复杂引擎的新实现成为可能。我们不知道最初的实现（生命的起源）是不是比后来的那些要困难得多，因为我们不是很清楚到底发生了什么。但显然讨论的所有系统，除了生命之外，它们的起源都是基于之前已经在运作的某种复杂引擎所发展出的结构才成为可能。这引出了一个问题，就是未来是否还会产生这个引擎的新的实现。如果是这样，将很有可能是从现存的复杂引擎的活动中涌现出来，并且有新的将信息作为指令进行编码和利用的方式。很难做出更具体的预测。谁能在100年前预测到社会（例如技术）的进化会导致解决困难的工程问题的进化算法的出现呢？现在进化算法已经被用于设计更强大的计算机。这一点有很深远的意义，下一章将进行探讨。

表**13.15**　种进化系统概要

进化系统	编码信息的本质	编码的形式
生命	生物化学反应的催化和调控以及大分子之间的特异性相互作用	4 种单体组成的 DNA 和 RNA 序列
适应性免疫	与特定微生物结合的蛋白质结构	DNA 核苷酸序列
学习和思维	基于感知输入的外部世界表示以及在这个世界中可能的将来行为	神经元之间的时空交流 模式
社会	向他人学习的知识以及人类行为的后果	人类语言模式
进化算法	计算机的执行动作和结果	计算机编码

进行中的计算意味着什么？

计算视角的一个自然推论是，你、你的身体、你的思维、你的成就以及你与他人的互动都是正在进行的计算的当前状态。如果这让你有些不安，这是因为你还没有完全接受计算无处不在的观念。计算与物理和生物过程的分离是过去400年科学认识发展过程中的历史性事件，本质上，不存在什么分离。

因为如果没有指令说明细节，大部分复杂事物都基本不可能存在，而我们所关注的大部分事物都是复杂的，我们的世界中的大部分事物也都依赖于指令才能存在。复杂引擎是简单却又强大的计算策略，在不知道怎样变化能改进指令的情况下也能对指令进行改进。反复对随机尝试进行选择就能做到。在某种编码媒介上记录这些选择就能不断对选择进行累计。这种方法使得复杂引擎成为可能，也解释了包括我们在内的本来不可能存在的各种复杂事物的存在。基于系统本身的细节及其所处的环境，复杂引擎能不断革新，似乎没有尽头。两个很明显的例子是生命——计算已经持续了35亿年——和人类社会的进化——已经持续了可能有20万年。在另一些实现中，引擎会达到一个目标然后停止，例如适应性免疫和大部分进化算法。人类思维的计算则贯穿人的一生（假设思维的确是依靠复杂引擎）。

我们习惯于将计算视为根据给定规则操作信息，产生输出然后停止的过程，输出可以是比如计算结果或打印的文档。复杂引擎的计算也产生输出，但这些输出在下一轮计算中又被用于输入。循环的过程可以无穷无尽，只有达到了预定的目标或者计算的物理基础崩溃才停止。正在进行的概率计算的中间阶段一个让人欣慰的特性是，虽然未来是未知的，但一切皆有可能。

第14章　未来

通往何方？

未来学家在预测未来方面有着糟糕的记录。这个可以解释。复杂引擎的核心特征就是在可能性空间穿行的轨迹是混沌的，无法从细节上预测它们将通往何方。创造性的源头是随机变化；不过预测未来仍然是一件有意思的事情。如果系统存在最优解，复杂引擎的变化就会朝着这个完美的答案前进；但即使是这种系统，通往可预测结果的路径仍然是无法预测的。回想一下第6章讨论的最多1问题：如果输出全为1，程序就会停止，但通往这个结果的路径有很多。选择作用于可选项，而实际提供的可选项都是随机的。随机变化提供可能性，选择做决定，未来取决于所选择的。如果系统没有最优解，比如存在协同进化，则结果会一直变化。同样，如果变化是随机的，则路径无法预测。

虽然很难，但有时候还是可以看出总体趋势。生物中就有这种收敛式进化。我们相当有把握认为在越来越干旱的气候下，植物会进化出刺和小叶子，而在越来越冷的气候中，哺乳动物会进化出更厚的皮毛。在特定条件下，进化系统预计会产生出越来越复杂的产物，对此有更多争议。在协同进化系统中，随着群体的进化，选择规则也会进化，因为各成员同时也是其他成员的环境。就像第12章讨论的，如果快速并且有计算需求的响应会带来优势，复杂性的一个方面——深度——就会增加。另一种很重要的进化是技术的发展，在技术的发展中，过去的成就为新的进展提供基础。当变化是基于之前的变化，效果将很惊人。一个例子是摩尔定律[1]，说的是过去半个世纪中，计算机的性能呈指数增长。

计算机学家、未来学家和科普作家雷伊·克兹维尔阐释了技术能力的指数增长和成本的指数下降的长远影响[2]。他重点关注3项技术，计算机技术、生物技术和纳米技术，他认为这些技术在不久的将来会融合。他的一些推论虽然让人难以相信，却也很难辩驳。下面我们逐个来探讨。

计算机技术

戈登·摩尔是集成电路的先驱之一，也是英特尔公司的共同创始人。1965年他提出了一个著名的预测，在今后的数十年里，集成电路中晶体管的数量每两年会翻一倍。图14.1绘制了各年生产的各种商用微处理器的晶体管数量。与预测惊人地相符。一些专家认为根据当前的技术水平，这种趋势在下一个十年很有可能还会持续。这意味着还会增长数十倍。

克兹维尔提出，摩尔定律是一个更久远趋势的最新发展，这个趋势包括5种计算技术，他称之为范式。5种计算技术分别是：机电设备、继电器、真空管、晶体管和集成电路。每种新范式的引入都使得设备变得更小，更快，也更便宜。图14.2是从1900年以来基于这5种范式的商用计算设备每千美元每秒的计算量。

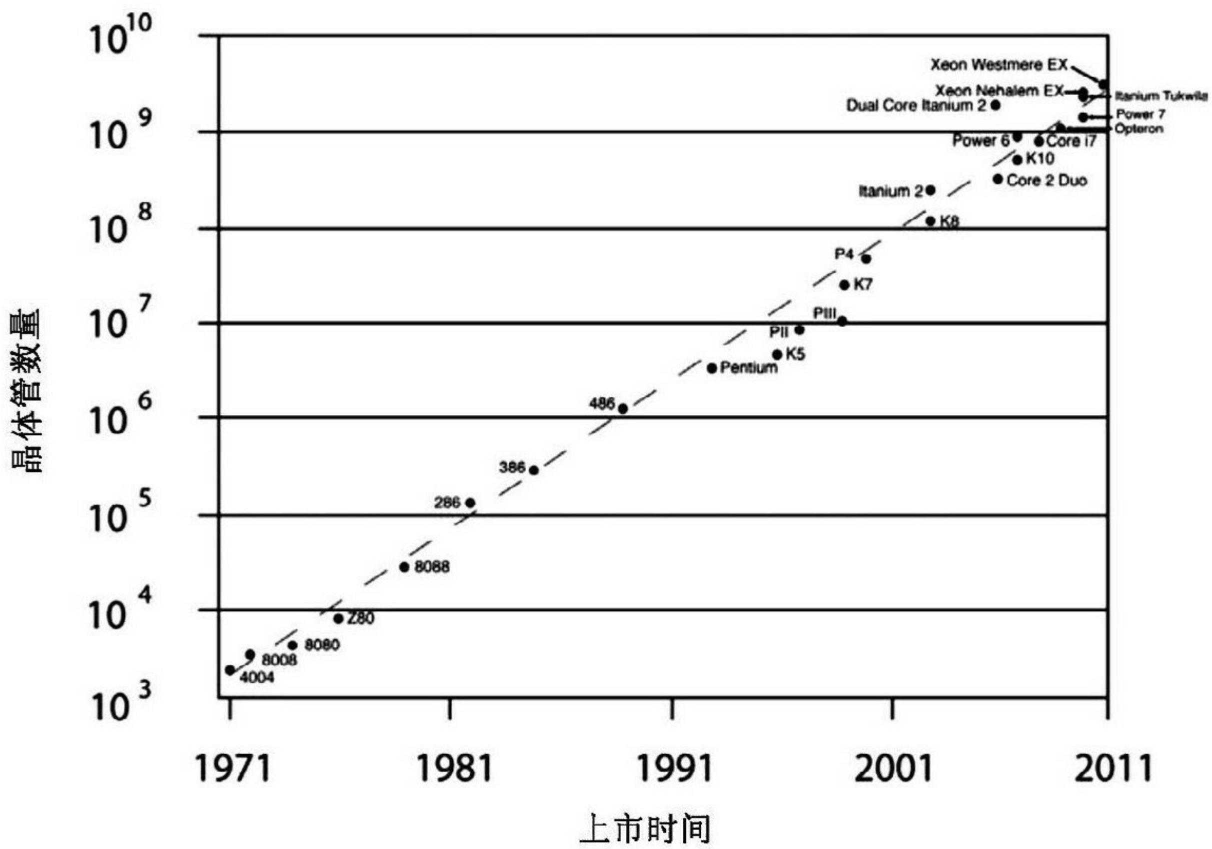


图14.1　CPU晶体管数量，1971-2011。横轴表示各种成功的商用芯片的上市时间，纵轴是每个集成电路的晶体管数量的对数。虚线表示每两年翻倍的预测，从1971年每颗芯片2　000晶体管开始。数据引自维基百科的摩尔定律词条

克兹维尔预测目前的范式——集成电路——在2020年之前会被新的范式——可能是分子计算——取代，从而继续推进这个趋势。他解释说，技术发展、成熟、然后被其他技术取代并不鲜见。他认为每种新技术都有3个发展阶段：

缓慢发展（指数增长的早期阶段）

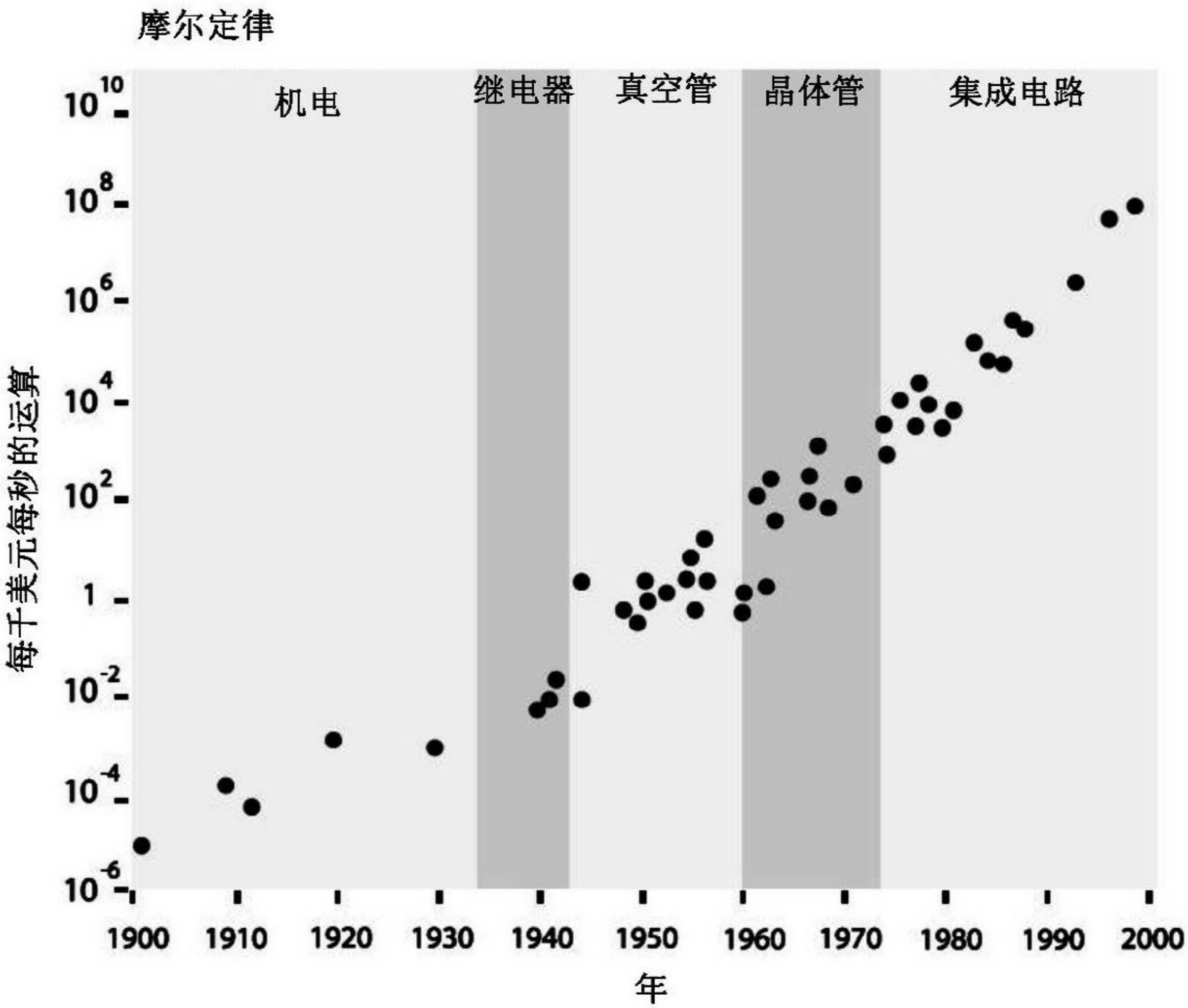


图14.2　从1900年以来各种商用计算设备每千美元每秒的计算量的对数。这些设备是基于5种不同的技术

快速发展（指数增长的爆炸式发展阶段）

范式成熟后的平稳阶段

实现最初的突破之后，任何新技术的发展都有一段缓慢发展的阶段，然后随着技术的进一步提升和越来越多工程师将技术应用到越来越多的场合，发展速度加快。这个阶段同时还伴随着成本的快速降低。最终，随着报酬递减效应，改进越来越难，技术的变化越来越小。这就是成熟阶段。技术进化成熟阶段的例子包括家用冰箱、微波炉、汽车化油器和铅酸蓄电池。如果技术性能可以度量，性能与时间的关系将是S曲线。当技术开始成熟，往往会被另一种新技术取代，新技术的发展同样

为S曲线。从而有可能形成技术的链条。图14.3（左图）体现了这个思想；而绘制成性能的对数与时间的关系时，结果（接近）为一条斜往上的直线（右图）。真实数据会有些波动。如果绘制的不是性能而是每单位性能的成本，则结果也是类似的，不过趋势是向下。

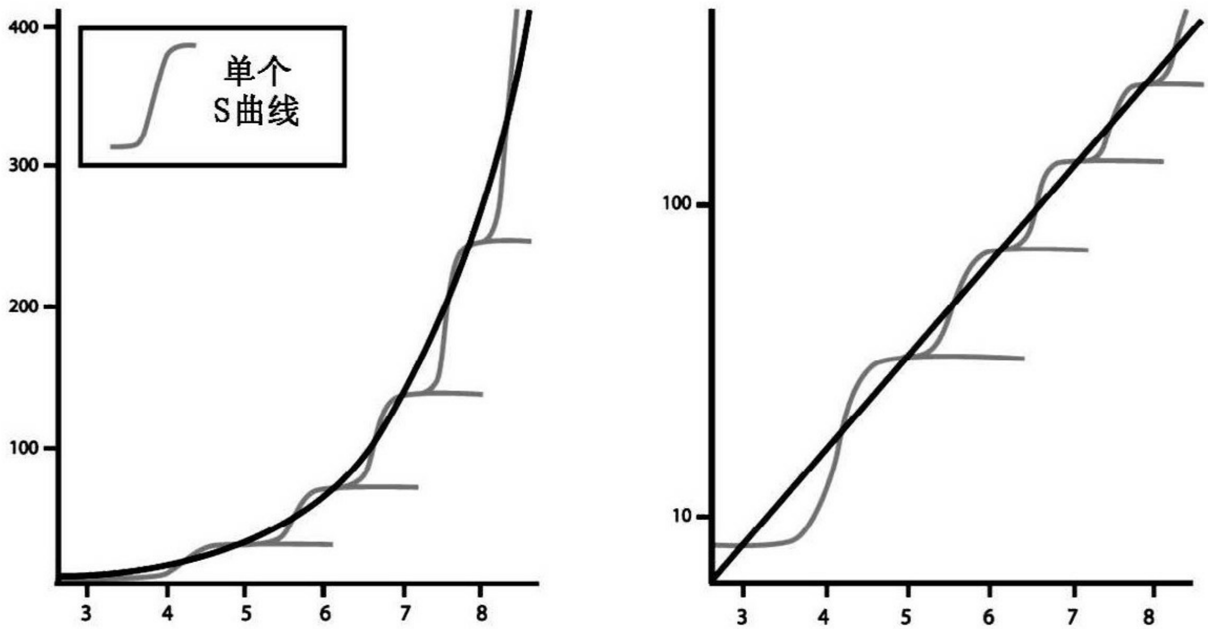


图14.3　技术进步似乎是由一系列技术相继出现构成，每种技术都有呈S形曲线的3个阶段。左边是性能与时间的关系，右边是性能的对数与时间的关系，表现为一条直线

计算机技术现在处于模仿人脑的起跑线上。人脑的逻辑结构完全不同于计算机，在某些方面更擅长。近年来高速计算机的并行化可能是通往这个方向的一步，因为大脑结构就是大规模的并行。大脑具有创造性，很擅长模式识别，也很擅长处理逻辑的不一致性，而这些都是数字计算机的弱点。为了实现这样的能力，需要对人脑的计算逻辑有更好的认识。

生物技术

摩尔定律的有效性无可否认，但是不是纯属巧合？其他技术是不是也存在长期的指数发展？克兹维尔关注的两项技术是生物技术和纳米技术。生物技术中最受关注的是改变和移植生物DNA的一系列能力。这些技术的核心是精确分析DNA序列的廉价技术，DNA测序技术目前已处

于快速发展阶段。这项技术最早于20世纪70年代发明，当时非常繁琐而且昂贵。到1990年，化学技术的发展以及自动化水平的提高使得人类基因组计划成为可能。这个由政府资助的计划有许多部门参与，目标是在15年内对人类DNA的30亿对碱基进行测序。在计划发起的时候DNA的测序成本大约为每对碱基10美元。当时估计随着技术进步测序会更快和更便宜。因此，计划的预算被设定为30亿美元，约为每对1美元，10倍的提升。计划最后提前了3年完成，花费了大约27亿美元。由于技术的发展极为成功，到2000年，由克雷格·文特尔领导的另一个团队采用最新技术只用了9个月花1亿美元就完成了第二次人类基因组测序（每对碱基3美分）。2007年詹姆士·沃森（DNA结构的发现者之一）宣布他获得了自己的DNA序列。完成这项工作的公司据说花了100万美元和2个月。也就是说，到2007年（大规模）测序的成本已经降到了每对碱基0.03美分。2008年美国国家卫生研究院宣布了1000组基因计划，准备测序1000组基因，每组的成本5万美元。2011年10月26日，X奖金基金会宣布了王者基因组X奖金。1000万美元奖金将奖励给第一个在30天内测序百名百岁老人基因组或最先做到每百万对碱基错误率低于1对并且每组基因成本低于1000美元的团队。竞赛从2013年1月3日开始。在竞赛开始的时候给各团队发放DNA样本。如果有人能获得奖金，实现了每组人类基因测序成本1000美元的目标（这很有可能），则只用20年测序一个人的DNA的成本就从30亿美元降到了1000美元，所需时间从15年降到了7小时。显然DNA测序技术的发展处于爆炸阶段。

这些进展是基于5条不同的技术发展路线：化学、自动化、并行化、微型化，以及分析巨量数据的计算技术的改进。自动化和并行化依赖于微型化，而所有这些，除了化学之外，又都依赖于计算机控制和分析。这体现了各领域技术的相互依赖和相互促进。

生物技术其他方面的进展很难度量，但毫无疑问许多前沿领域都在以惊人的速度发展。这部分是因为廉价DNA测序数据的迅速增长，部分是因为强大的计算技术的迅速发展。

这对未来意味着什么？一是针对个人的医疗。美国国家卫生研究院之所以将纳税人的钱用于降低测序的成本和提高速度，就是为了将DNA序列纳入每个人的医疗数据。这样就有可能针对每个人的基因组量身定制医疗方案。目前，由于每个人对药物的反应因人而异，因此药物的作用很受制约。药物与蛋白质相互作用，而身体里的蛋白质又由DNA序列

决定，因此如果知道了DNA序列，就能收集与每个病人的蛋白质有关的药物反应的数据。

另一个还很遥远但并非遥不可及的可能性是设计生物。如果生物学的发现以及相关的技术继续加速发展，不难想象有一天基因工程师会坐在电脑前“设计”新种玉米或清除油污的新微生物之类。我们目前已经处于重新设计病毒用于医疗用途的前沿。

重新设计病毒的目的是基因治疗。一些人天生有基因缺陷。有人提出将正常基因移植到体细胞中来进行修复。病毒可以携带基因进入体细胞。这个想法已经开始了临床试验，目前还在改进，10年之内有可能成为常规疗法。

纳米技术

克兹维尔技术未来观的第3部分是纳米技术。这个领域关注的是极度微型化。设备尺寸的缩减由来已久。19世纪末，工业技术由巨大的机器主导，用巨大的齿轮和杠杆完成各种工作。其中一些大型机器仍在服役，用于升起吊桥或关闭水闸，但大部分已经被更小而又更强大，更能胜任那些工作的机器替代。举个例子说明进步有多大，2003年加利福利亚大学伯克利分校的一个团队宣布制造出了比人的头发直径小300倍的马达[3]。

由于设备的范围很广，很难用图表说明设备尺寸随时间减小的趋势。不过可以以钟表为例。埃及人在公元前14世纪左右发明了水钟。最初只是底部有小洞的陶碗。在内侧刻线，就可以用水的流失表示时间的流逝。后来改进为用流水驱动运动的部件。14世纪前后欧洲发明了落水钟。最初的设计并不精确，但可以四处移动而无需调整铅垂。第一块怀表出现在1462年左右，有16世纪初制造的怀表留存至今。这些早期的设计用今天标准来看很粗糙。第一块腕表出现在1862年，到20世纪20年代商品腕表已经能够做得很小，足以吸引具有时尚意识的女士。1909年亨利·瓦朗获得了第一台电子钟的专利；这些钟被安装在漂亮的家具中。2008年爱普生东洋通信发布了一款尺寸只有3.4毫米×1.7毫米×1毫米的精准时钟芯片。当然这款芯片没有钟面，否则你要用放大镜才能看时间。在技术的发展史中，钟表变得更小、更便宜、也更精确。大部分技术的

发展都有类似的经历。

我们之前讨论了晶体管。晶体管尺寸的缩小也符合摩尔定律。由原子组成事物的微型化必然具有极限。一个物体不可能比单个原子还小，而晶体管、马达或钟表等设备由多个部件组成，每个部件都不可能比单个原子更小。微型化的极限是操作单个原子。1990年，IBM的科学家朝这个方向迈出了惊人的一步，他们能用原子力显微镜在平整表面排列单个原子。图14.4展示了用35个原子排列的字母IBM [4] 。从排列氙原子到制造出原子级别的马达或者能修复分子尺度损伤的纳米机器人还有很长的路要走，但许多有远见的科学家和工程师都将此视为终极目标。

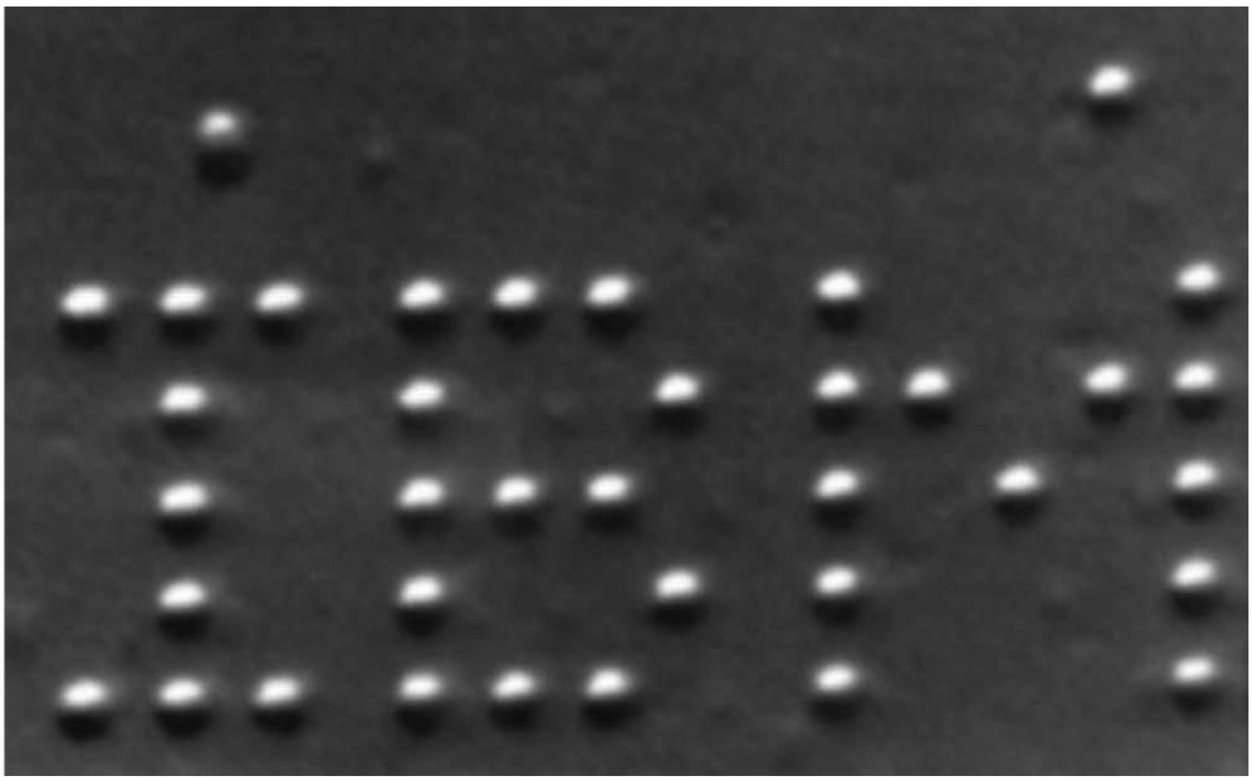


图14.4　在镍金属表面用氙原子拼出的字母IBM（自然出版集团拥有版权）

这个领域的许多学者都从生物学获得灵感。活细胞中就存在许多微型机器，包括马达、棘轮、泵、光传感器、微小的运输系统，甚至时钟。它们都非常小。一种制造微型分子设备的策略就是向生物学习。因此纳米技术的一些方面也将与生物技术融合。在我们有生之年能不能看到注入血管识别和摧毁癌细胞的纳米机器问世？有可能。

制造微型设备需要计算，因为微型设备需要有很高的装配精度，而

且原子和分子的行为很难预测和控制。显然这是相互依赖，生物技术和纳米技术的发展依赖于越来越高的计算能力，而计算能力的发展反过来又依赖于纳米和生物技术。

许多人共同关心的一个问题是，随着技术的发展，机器会不会有一天取代我们？克兹维尔则从另一个角度看问题。他认为我们将会与技术融合！毕竟我们的大脑就是信息处理设备。大脑是生物计算机，与电脑具有完全不同的逻辑结构。有3个指标说明了这方面的趋势：以“人机交互”为主题的毕业课题数量（在美国2000年1个也没有，到2009年有69个），亚马逊网站关于脑机交互的新书数量（2007年和2008年达到6本），以及关于这个主题的科普文章和新闻的数量（很多）。2008年11月2日，在CBS的节目中，斯科特·佩利说道：

在许多安静的实验室中，正在发展一项惊人的技术，将人脑直接连接到电脑。这像是一次突然的人类进化跨越——这次跨越将在未来让瘫痪的人再次行走，让截肢者借助仿生肢体行动。

故事主角斯科特·麦克勒患有肌萎缩性（脊髓）侧索硬化，这是一种神经退化疾病，让他除了眼睛之外全身瘫痪。通过类似头盔的可穿戴设备，斯科特能够参加电视访谈，通过分析他想说的话，计算机就能替他发声。现在这项技术处理的速度还很慢，斯科特回答问题时平均每个字母要花20秒。但还是能让他交流，而且交流得相当好。2008年艾莫提夫系统公司发布了用于计算机游戏的第一个商用脑机接口。这是一种头戴式设备，有14个传感器，穿戴者可以通过意念控制角色的行为。零售价299美元。目前这还是新鲜事物，因为大多数人还是更习惯用手，但谁又知道十年或二十年之后会怎么样呢？

另一项更成熟的技术是耳蜗埋植。2009年有15万深度耳聋者依靠这种设备恢复了听力。其原理是将小型麦克风连接到音频处理器，通过植入听神经的电极阵列输出电脉冲。然后大脑会将这些脉冲传换为听觉。植入的耳蜗替代了内耳的功能，让患者可以恢复听力，当然目前的效果还比不上健康的耳朵[5]。

目前脑机接口还比不上正常的身体功能，但这项技术的发展还不到15年，正在迅速提升。也许有一天你穿戴的设备可以让你的大脑像计算机一样快速而且精确，同时又还保留了人类的优势。这一天并非遥不可

及。

克兹维尔本身是计算机专家。他注意到现在的计算机能力的指数增长和成本的指数递减只进行了40年，我们生活的世界就已经彻底改变。如果光速对信息传递速度的限制可以被绕开，那么只需200年整个宇宙都会改观。这纯粹是科幻，但在不远的将来我们将看到两个重要改变。一是将会出现由计算机设计的计算机。这将使得技术脱离人类思维的限制。现在，计算机芯片的设计已经非常复杂，人类工程师已无法完全掌握；软件在英特尔新一代芯片的设计中扮演着重要角色。人类工程师还要多久才会退出这个循环？二是计算机将能够模拟人类的思维过程。这将使得计算机可以与人类互动。

如果脑机接口能让人类和机器交换大量比特信息，人类思维将脱离人体的限制。由于信息可以复制，因此不必限制在特定的物理系统中。如果我们的思维可以被复制给具有人类思维能力的机器，我们的思维就不会随我们的身体死去，我们（我们的记忆和思维）将变得不朽。克兹维尔认为这将在21世纪40年代实现，并且在力争让自己活到那个时候。不清楚在他的想法中是不是要让自己的身体也一直活下去，不过我可以肯定，如果问他，他不会在乎这一点！

我们准备好了参与复杂引擎的新的实现吗？

在思考我们星球的历史上复杂引擎的新的实现时，会注意到两个模式。首先，实现了这个引擎的系统似乎会引发新的实现（第13章曾经讨论）。另外，新的实现之间的时间间隔越来越短。生命起源在某种程度上是所有实现的基础。我们还没有发现哪种复杂引擎与生命无关。图14.5表现了这种趋势。图中略去了适应性免疫系统，因为它似乎不太可能产生对人类未来很重要的变化；而且将动物思维包括了进来，虽然还不确定复杂引擎是否在其中具有核心地位。

故事是这样：40亿年前出现了生命（事件1）；大约2亿年前出现了具有初级思维能力的动物脑（事件2）；大约20万年前开始出现人类语言和基于语言交流的社会（事件3）；60年前出现了基于计算机的进化算法（事件4）。图14.5画出了各起源时间的对数值。图中只有4个点，但趋势很惊人，如果继续下去，预计下一次起源随时会出现！下一步会

不会是计算机设计计算机？会不会是计算机的体系结构将复杂引擎包括进去？会不会是人类思维与计算机硬件的融合？会不会都发生？还是其他事件？时间会证明一切，如果我们没有消灭自己，很有可能在不远的将来看到复杂引擎带来新的意外[6]。这将是让人兴奋的时刻。

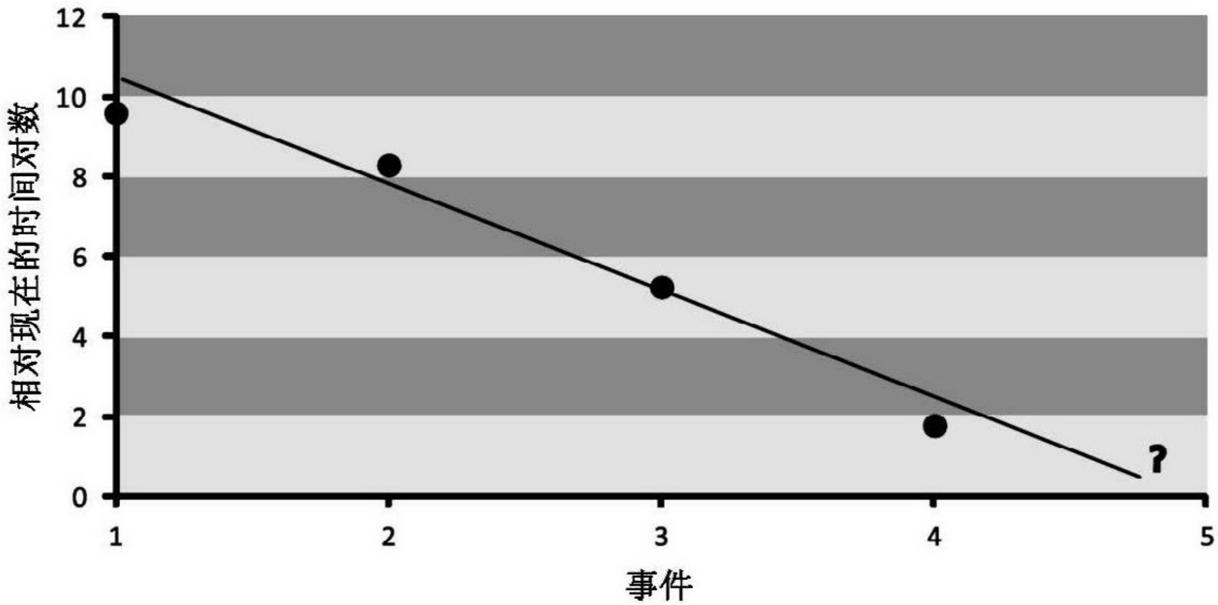


图14.5　事件发生时间的对数图。数字表示的事件为：（1）生命的起源，（2）思维的起源，（3）基于语言的人类社会的起源，（4）进化算法的起源，（5）现在

致谢

首先要感谢我的同事丹·阿什洛克和杰克·卢茨，没有他们持之以恒的兴趣和耐心，我不可能深入理解进化计算和理论计算机科学。没有他们富有启发的讨论就不可能有这本书。这本书的第一个读者是马特·帕蒂兹，他当时是爱荷华州立大学的研究生，现在是阿肯色大学计算机科学和工程系的助教，他阅读了书稿，并且给出了正面的反馈。还有许多同事，包括利·特斯特法新、阿兰·阿瑟利、詹姆斯·莱思罗普、詹姆斯·莫罗、伊桑·斯科特、唐·坂口、道格·芬尼莫尔、史蒂芬·卡瓦莱尔和加里·霍华德，都给出了许多中肯的意见和建议。两位系主任杜安·恩格尔和马丁·斯波尔丁也抽出宝贵时间阅读了书稿。还有数十位大学生也阅读了书稿并给出了意见，在此不再一一致谢，你们的付出对我很重要，我希望你们能喜欢这本书。学生们还帮我指出了不易理解的段落。他们的许多建议都被采纳。一些章节所探讨的思想最初是在我教授的复杂适应系统的课堂上拿出来讨论的。其中大部分图表是由爱荷华州立大学学习生物和医学制图的学生绘制的。他们是埃米·迪坎普、奥德丽·吉福德、米根·赫瑟尔顿、布兰登·霍尔特、塔莎·奥布林、梅格·皮珀布林克和肖恩·西贝尔斯基。他们的酬劳是由爱荷华州立大学出版资助基金提供的。另外我还要感谢我的出版商，彼得·普雷斯科特，是他帮助我编辑和走完复杂的出版流程；同时还要感谢哥伦比亚大学出版社的帕特里克·菲茨杰拉德和布里奇特·弗兰纳里—麦科伊整理书稿并制版。最后感谢我的妻子贝齐、女儿凯特和玛吉，以及女婿戈登·鲍恩和伊桑·斯科特，一直鼓励我的写作并提出意见和建议。尤其要感谢他们让我意识到我所阐释的概念并不像我想的那样简单；尤其感谢贝齐和凯特，她们本身就是很有造诣的作家。

注释

引言

[1] Michael Ruse在Darwin and Design（2003）一书中给出了关于目的复杂性的哲学思想的简明历史。

[2] Information and the Nature of Reality（2010），Paul Davies和Niels Henrik Gregersen编撰，收录了许多知名学者的综述文章。这本书是目前哲学研究的前沿资料，涵盖物理学和宗教。塞思·劳埃德的章节“计算的宇宙”概括了他之前的Programming the Universe（2006）一书的主要思想。

[3] 我认为理查德·道金斯的《盲眼钟表匠》（1986）是为非专业读者阐释生物进化的最好读本，虽然计算机模拟有些过时。优秀的大学新生会喜欢这本书，基础差一些的则有些困难。对这部分学生，可以读道金斯后来的一些书，例如Climbing Mount Improbable和The Ancestors Tale：A Pilgrimage to the Dawn of Evolution。

[4] Daniel W. McShea和Robert N.Brandon2010年的新书Biology’s First Law, The Tendency for Diversity&Complexity to Increase in Evolutionary Systems探讨了标准的达尔文学说的一半：选择之前的可继承多样性的产生。如果抛开选择，多样性增加就是多样性产生的自然结果，作者将此提升为了“定律”。之前还有人对选择也做过同样的尝试，提出了“自然选择律”。这些尝试能否进入进化的教科书还有待观察。

[5] 原因在于“定律”一般表示为简单的命题或数学方程。进化是一个过程，进化计算的输入往往很长而且复杂，产生的结果也很复杂。简单的命题或关系很少适用于这种情况。

[6] 对于对进化感兴趣的读者，除了道金斯，还有一个要推荐的作者是丹尼尔·丹尼特。丹尼特的Darwin’s Dangerous Idea（1995）可能是与这本书最接近的。他是一流的哲学家，对进化理解深刻，并且擅长以易于理解的方式写作。

[7] 托马斯·库恩（《科学革命的结构》，1962）被公认提出了“范式转变”作为科学进步的基本特征。库恩和卡尔·波普尔是20世纪两位最有影响力的科学哲学作家。波普尔认为科学进步具有进化特征。

[8] 史蒂芬·古尔德在Wonderful Life：The Burgess Shale and the Nature of History（1989）一书中有力地阐明了这一点。这是一本有趣而且富有洞察力的书，介绍了生活在“寒武纪大爆发”时期的生物。这是动物化石第一次出现在地质记录中的时期。

[9] 赛斯·劳埃德是理论物理学家，与圣塔菲研究所关系密切。他最近的书Programming the Universe：A Quantum Computer Scientist Takes on the Cosmos（2006）阐释了信息在宇宙学中的核心地位。

[10] 如果对心理学和大脑感兴趣可以读Steven Pinker的书。How The Mind Works（1997）是我的最爱。Pinker让人信服地论述了大脑是信息处理器官以及我们的行为都是其活动的产物的观点。

[11] 社会进化一直是极具争议的领域，很难为读者提建议。我推荐3本书：道金斯的《自私的基因》（1976），David Sloan Wilson的Darwin’s Cathedral（2002），以及Eric Beinhocker的The Origin of Wealth（2006）。各有特色：道金斯写遗传学，David Sloan Wilson（从群体生物学家的视角）写宗教，Beinhocker写经济。他们一起勾勒了进化活动在所有社会组织中无所不在的地位。

第1章

[1] 这个食谱来自HGTV网站：http：//www.hgtv.com/hgtv/ah_recipes_poultry_eggs/article/0，1801，HGTV_3189_1394919，00.html。在网上搜“简易蛋奶酥”可以找到很多类似食谱。

[2] Great Principles of Computing：http：//cs. gmu.edu/cne/pjd/GP。如果想对计算有更深刻的认识，这个网站很不错。

[3] 在信息科学中意义是一个很难界定的词，部分是因为意义是相对的。取决于背景。人们可以量化地说一个信息与另一个信息的相关程度，或者导致了另一个系统多大变化，或者一个进化系统获取了多少关

于其环境的信息（Adami et al.2000）；但人们不能量化一定量的信息对某人或某物的意义。

[4] 以下式子可以计算系统的香农熵：$Æ=-\sum p_i \log_2(p_i)$。符号∑表示对n种可能状态求和，p_i 是第i种状态的概率，$\log_2$ 表示求以2为底的对数。Æ的单位为比特。这个公式用起来很简单。假设系统有2种可能的状态，则公式变成$Æ=-[p_1 \log_2(p_1)+p_2 \log(p_2)]$，其中$p_1$ 为状态一的概率，p_2 为状态二的概率。如果两个状态的概率相等，则$p_1 =1/2$，$p_2 =1/2$，$1/2\times\log_2(1/2)+1/2\times\log_2(1/2)=-1$，因此Æ=1比特。只有两个等概率状态的系统携带的信息量为1比特。如果概率不相等，则Æ的值要低一些。香农将Æ称为熵，但他和其他人都将Æ视为对信息的度量。当系统所有状态的概率（p_i 项）相等时，Æ达到最大值。

[5] 有一个简单的办法可以帮助理解为什么对象携带的信息（不考虑意义）等于以对象为具体状态的系统的熵，就是将对象的信息定义为对象的熵与系统的熵的差值。由于单个对象的熵为0（概率为1，而1的对数为0），因此刻画对象的信息就等于系统的熵。

[6] 要想深入了解这些概念必须学些数学。有一本被广泛采用的教材是科弗（Cover）和托马斯（Thomas）著的《信息论基础》（Elements of Information Theory，1991）。

[7] 这个思想扩展了注释5中给出的逻辑。当（注释4中）香农公式所有的概率相等时，系统达到最大熵。这等同于消息中的所有位随机选择。如果概率不等，则序列就在某种程度上不那么随机，而非随机意味着具有结构。Adami et al.（2000）用进化系统的实际熵与最大熵之间的差值定义复杂性，并用来度量系统从环境中获得的信息。其中的思想是，如果进化的位串是随机的，则没有携带环境信息，而如果概率不等，则表明获得了信息。我不认为复杂性一词用在这里很适当，作者还不如用负熵，与热力学意义保持一致。

[8] 所有关于进化的书都有这样一张表。细节可能不同，但都可以扩展或归纳成这5条。

[9] 《攀登不可能之山》（Climbing Mount Improbable）是《盲眼钟表匠》的续篇。道金斯在这本书中想澄清一些前一本书读者误解或不理

解的地方。因此这本书比《盲眼钟表匠》要容易读些，但如果想更好地理解进化，建议两本都读。

[10] Dennis Bray的Wetware, a Computer in Every Living Cell（2009）探讨了细胞生命行为构成计算的议题。

[11] 牛顿引力定律的数学形式如下：$F=G\cdot m_1 \cdot m_2 /r^2$ 。F是质量为m_1和m_2 ，距离为r的两个物体之间的引力，G为常数。后面我会强调这个以及类似的公式只是简单关系。如果你对此感到怀疑，比较一下引力定律和蛋奶酥的配方或数千行代码组成的程序。

[12] 道金斯、丹尼特、平克（Pinker）、埃德尔曼（Edelman）和拜因霍克尔（Beinhocker）的著作支持了这种论断。关于这个主题还有一本有趣的书是乔治·戴森（George B.Dyson）的《机器中的达尔文》（Darwin Among the Machines，1997）。

[13] Michael Behe的Darwin’s Black Box（1996）可能是论证无法还原的复杂性的最具影响的著作。这类论证都有一个问题，我称之为缺乏想象力，更常见的说法是“基于个人的不相信的论证”（道金斯，《盲眼钟表匠》，p.38）。换句话说，不经证明就做出“这根本不可能发生”这样的论断。随便到网上搜索一下Behe就能找到许多文章解释为什么他的论证站不住脚。对这本书Robert Dorit在1997年9-10月的《科学美国人》发表了一篇很有说服力的书评。

[14] Sean Carroll的Endless Forms Most Beautiful：the New Science of Evo Devo and the Making of the Animal Kingdom（2005）。在我看来，这本书是向普通读者介绍通过进化创造动物身体结构背后的分子机制的最好读本。

[15] 这种图被称为分形，图中例子被称为Julia集，以纪念Gaston Julia第一个研究了它们的性质。图1.1是Matefitness项目（Genoa, Italy, www.matefitness.it）的成果，经许可引用。

[16] 丹尼尔·丹尼特在Freedom Evolves（2003）一书中详细解释了意向性立场的含义。

[17] 道金斯在《盲眼钟表匠》中直接引用了William Paley，然后给出了自己的论证，指出曾让达尔文印象深刻的Paley的论证现在被发现是不成立的。

[18] 加里·齐科（Gary Cziko）在《没有奇迹》（Without Miracles，1995）一书中提出所有复杂适应的例子都源自选择。他的结论源自唐纳德·坎佩尔（Donald Campbell）和卡尔·波普尔（Karl Popper）的著作，他在书中评论了他们的观点。

第2章

[1] 计算大原理网站上（http：//cs.gmu.edu/cne/pjd/GP）有很多关于计算的思想，很富有启发。

[2] 对于想深入了解进化的读者，我极力推荐丹尼尔·丹尼特的《达尔文的危险思想》（1995）。丹尼特的研究领域是哲学。

[3] 关于逻辑门及其电路设计在网上可以找到很多内容。搜索“逻辑门”就能找到很多信息，远比我在这里给出的多。

[4] 斯蒂芬·沃尔弗拉姆（Stephen Wolfram），《新科学》（A New Kind of Science，2002）。这本书有1000多页，但是图很多，容易阅读。沃尔弗拉姆希望他书中的思想“不仅能影响科学和技术，同时也能影响到一般性思维。复制图2.3-2.6需经Wolfram Research公司同意。

[5] 同上。

[6] 埃里克·温弗里的网页位于http：//www.dna.caltech.edu/～winfree/。他和他的合作者研究DNA模块的编程结构。

[7] 图2.7和2.9的思想源自“Algorithmic Self-Assembly of DNA Sierpinski Triangles，”Paul W.K.Rothemund, Nick Papadakis&Erik Winfree, PLOS Biology, pp.2041—2052，Vol.2，December 2004。图2.8是直接引自该文。经创作共用许可契约（Creative Commons License Deed, http：//creativecommons.org/licenses/by/2.5/）许可使用。

[8] 杰克·卢茨（Jack Lutz）是爱荷华州立大学的计算机学教授，一

直给研究生教授拼砌理论。这句话引自他2007年在课堂上说的话。

第3章

[1] 如果对宇宙起源和“万有理论”感兴趣，格林（Brian Greene）的《宇宙的琴弦》（The Elegant Universe，2003），塞思·劳埃德（Seth Lloyd）的Programming the Universe（2006），李·斯莫林（Lee Smolin）的The Life of the Cosmos（1997），以及保罗·戴维斯（Paul Davies）的《上帝与新物理学》和《实在终极之问》等一系列书都是很好的读物。

[2] 参见沃尔弗拉姆的《新科学》中的数百个例子。

[3] 这一章的标题是受斯图尔特·考夫曼（Stuart Kauffman）在《宇宙为家》（At Home in the Universe，1995）一书中提出的“免费的结构”一语的启发。考夫曼用这个词描述非平衡态条件下自发形成的结构。这个词贴切地形容了所有不需要指令就能形成的结构。

[4] 库仑定律的数学表达式为，

$$F = -\frac{k \times q_1 \times q_2}{r^2}$$

式中F是力，k是常数，q_1 和q_2 是两个电荷大小，r是电荷之间的距离。通常电荷为正q就为正，因此如果另一个电荷为正一个电荷为负，F就为正（电荷吸引）；如果两者都为正或为负，F就为负（电荷排斥）。

[5] 含时薛定谔方程

$$ib\frac{\partial \psi}{\partial t} = -\frac{b^2 \partial^2}{2m\partial x^2} + V(x)\psi(x,t) \equiv H\psi(x,t)$$

是描述量子力学行为的基本方程。它描述了物理系统随时间变化的波动方程。Ψ是波函数，V（x）是势函数，H是被称为哈密顿算子的数学函数。这个方程的解就是所允许的能级。

[6] 牛顿万有引力定律：$F=g\cdot m_1 \cdot m_2 /r_2$ ，其中g是常数，m_1 和m_2 是两者质量，r是两者距离。

[7] 贝纳德流等类似现象的驱动力长期被误解。关于表面张力的核心作用是怎样被发现的历史，参见科施米德尔的Bénard Cells and Taylor Vortices（1993）。

[8] 幂律公式为$y=ax^k$ 。这个公式描述的是指数关系。这种关系在自然界很常见。等式两边取对数得到log（y）=k·log（x）+log（a）。以log（y）和log（x）为坐标轴画图，会得到斜率为k的直线。如果数据在对数-对数坐标上形成了直线，就表明存在幂律。

[9] 佩尔·巴克（Per Bak）的How Nature Works（1996）展示了自然界和试验中观察到的许多幂律的例子，并强调了自组织对于理解许多现象的重要性。

第4章

[1] 我认识内特·约翰逊时他是一名研究生，当时他在利用进化算法为中美洲的穷人改进简易燃木炉。他在大学阶段学习机械工程课程是在2003年。

[2] 威廉·佩利的《自然神学》（Natural Theology-or Evidences of the Existence and Attributes of the Deity Collected from the Appearances of Nature，1802）是最著名的认为有更高的力量设计了生命的论证，道金斯的《盲眼钟表匠》的标题受此启发。道金斯有力地解释了科学发现是如何否定了佩利的观点。

[3] 这一节省略了许多分子细节。每种密码子氨基酸对都有不同的tRNA，还有特定的酶确保正确的氨基酸会首先与每个tRNA连接。一个特殊的密码子AUG启动mRNA上的过程，3种终止密码子都可以让过程结束。过程还需要各种被称为翻译因子的“助手”蛋白质。想了解更详细内容的读者可以阅读任何一本大学水平的生物学入门教材中的分子遗传

学章节。我一直使用尼尔·坎贝尔（Neil Campbell）和简·里斯（Jane Reese）的《生物学》。这本书现在是第9版，最新版的第一作者现在是简·里斯。

[4] 网站上有T_4组装和感染过程的漂亮动画：http：//www.seyet.com/t4_academic.html。书中的图引自2003年的综述文章，P.G.Leiman, S.Kanamaru, V.V.Mesyanzhinof, F.Arisaka，&M.G.Rossmann, Structure and Morphogenesis of Bacteriophage T 4，Cellular and Molecular Life Sciences, vol 60，pp.2356—2370。

[5] 随便找一本入门教材，例如坎贝尔和里斯的《生物学》，都会有更详细的阐释。完整的阐释要参考更专业的教材。我推荐的两本是马克·比尔的《神经科学：探索大脑》（Mark F.Bear, BarryW.Conners，&Michael A Paradiso, Neuroscience：Exploring the Brain，3d ed.，2007）和戴尔·普维斯的《神经科学》（Dale Purvis, Neuroscience，5th ed.，2012）。

[6] 《自私的基因》（The Selfish Gene，1976）是理查德·道金斯的第一本书，在他写的生物和社会学书籍中可能是最有影响力的一本。其中尤其是有两点经受了时间的检验。第一，身体是传播基因的工具；第二，文化中传播的思想概念可以用基因的术语来理解。在书中他提出了拟子（meme）一词来描述文化中与（DNA）基因相对应的概念。

第5章

[1] 在《银河系搭车客指南》（1979）中，道格拉斯·亚当斯根据量子力学的多重宇宙解释，基于未来科技能实现跨平行宇宙旅行的设想，想象了幽默奇幻的星际之旅。

[2] “没有词汇的数字认知：来自亚马逊的证据。”彼得·戈登（Peter Gordon）2004年10月15日发表在《科学》杂志的文章提供了清晰的证据指出有些语言的计数不能超出2。

[3] 有很多很好的网站详细介绍了如何用罗马数字进行初等算术。搜索“罗马数字算术”。

[4] 丹尼尔·丹尼特在《达尔文的危险思想》中借鉴了豪尔赫·路易斯

·博尔赫斯（Jorge Luis Borges）在《迷宫》（Labyrinths，1962）中想象的所有可能书籍的图书馆，“巴别塔图书馆。”丹尼特想象了巴别塔图书馆的子集，称为孟德尔图书馆，图书馆中有所有可能的DNA序列。

[5] Small probability reference：Risk to groundlings of death due to airplane accidents：a risk communication tool. Bernard Goldstein, et al.，Risk Analysis 12（1991）：339—341。数据来自美国，1975-1985。

[6] 《物种起源》，查尔斯·达尔文，1859。这本书可能是有史以来影响最大和被引用最多的科学著作。到现在仍然有可读性。

[7] 肖恩·卡罗尔（Sean Carroll）的Endless Forms Most Beautiful（2005）可能是最好的介绍进化发育生物学的大众读物。

第6章

[1] 机器学习或人工智能是一门相当专业的学科，目标是让计算机能够学习。《科学美国人》（Scientific American）2002年出版的文集《理解人工智能》（Understanding Artificial Intelligence）是一本不错的介绍性读物。

[2] John Holland, Adaptation in Natural and Artificial Systems：An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence，2nd ed.，1992；David Goldberg, Genetic Algorithms in Search Optimization and Machine Learning，1989。这两本书都很适合非专业读者。

[3] 塔耳塔罗斯（Tartarus）是阿斯拓洛·特勒（Astro Teller）发明的计算机游戏，其祖父爱德华·特勒（Edward Teller）曾主导美国氢弹计划，外祖父杰拉德·德布鲁（Gerard Dedbreu）是诺贝尔经济学奖获得者。特勒说自己是企业家、科学家和作家。他在卡内基·梅隆大学读研究生时为研究机器学习发明了这个游戏。丹·阿什洛克（Dan Ashlock）吸收了这个思想，并与其学生一起深入研究，他2006出版的进化计算专著（Ashlock，2006）用一章篇幅对塔耳塔罗斯实验进行了探讨。

[4] Goldberg，1989；Ashlock，2006.

[5] 所有生物入门教材都用这种图，通常认为是休厄尔·赖特最先使用。Sewell Wright, The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution.Proceedings of the Sixth International Congress onGenetics, pp.355—366，1932。

第7章

[1] 关于发育过程中的细胞迁移和程序性细胞死亡在任何现代发育生物学课本中都有讨论，例如Lewis Wolpert, et al.，Principles of Development，3rd ed.（2006）和Scott F.Gilbert, Developmental Biology，8th ed.，（2006）。

[2] 杰拉尔德·埃德尔曼（Gerald Edelman）因免疫学的研究在1972年获得了诺贝尔生理学或医学奖，然后就转向了神经系统的研究。他在《神经达尔文主义》（Neural Darwinism，1987）一书中阐释了他关于变异和选择在大脑运作中起核心作用的思想。埃德尔曼与朱利奥·托诺尼（Giulio Tononi）合著的《意识的宇宙》（A Universe of Consciousness，2000）对他的思想进行了更浅显的阐释（译注：参见埃德尔曼的《比天空更宽广》，湖南科学技术出版社，2012和《第二自然：意识之谜》，湖南科学技术出版社，2010）。

[3] 最新的大学生物入门教材通常会用1章介绍哺乳动物免疫系统。阅读这样的课本章节是了解这个领域的不错途径。进一步阅读可以参考专门针对这方面的中级水平课本，例如Janeway, Shlomchk, Travers，&Walport, Immunobiology（2004）。

[4] 在人类医学中通常使用HLA（人类白细胞抗原）一词而不是MHC（主要组织相容性）。

第8章

[1] 同样，任何一本生物学入门教材都会介绍配对和重组（有时候也叫“交换”）。

[2] David Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning（1989）。戈德堡对遗传算法的定义中包含有重组。

[3] 达尔文，《物种起源》，1859。

第9章

[1] 斯图尔特·考夫曼，Origins of Order（1993），At Home in the Universe（1995，中译本《宇宙为家》，湖南科学技术出版社），Reinventing the Sacred（2008）。考夫曼是圣塔菲研究所（SFI）的长期成员，也是向大众介绍复杂性科学的新发现以及复杂适应系统最有影响力的人之一。

[2] 图9.8引自Cyran等人2003年发表在《细胞》期刊的一篇综述文章。细胞中生物钟的分子细节还没有完全清楚，各物种之间也有明显区别。在网上对此进行深入了解的一个好地方是NIH网站（http：//www.nih.gov/），或者直接搜索“生物钟”。

[3] 这个列表是直接引自维基百科上的条目。有两本可读性很强的书是约翰·霍兰的Hidden Order（1995，中译本《隐秩序》）和Emergence（1998，中译本《涌现》）。

第10章

[1] Eric Kandel, In Search of Memory, the Emergence of a New Science on Mind（2007）。埃里克·坎德尔回忆了他成功的神经科学家学术生涯，同时对现代思维科学进行了很好的综述。

[2] Karl Popper, Objective Knowledge：An evolutionary Approach，1972（中译本：波普尔，《客观知识：一个进化论的研究》，上海译文出版社），第六章，关于云和钟。这一章是1965年4月21日在华盛顿大学的亚瑟·霍利·康普顿纪念演讲的正式文本。这个精彩的讲座勾勒了人类思维中的随机性和选择的作用，同时也是对波普尔的工作的精彩论述。

[3] 对于肌肉功能的基本介绍，再次向读者推荐大学生物入门课本，例如坎贝尔和里斯的《生物学》（2008）。

[4] 要想进一步理解混沌，有一本不错的书，Robert L.Devaney, Chaos, Fractals, and Dynamics, Computer Experiments in Mathematics，

1990。

[5] Arthur H. Compton, The Human Meaning of Science, pp.45—46（1940）。康普顿从这个例子开始论证，由于量子力学，经典因果律在分子层面不再成立。据此康普顿认为，让人欣慰的是，人类的自由意志与现代物理学相容。

[6] 哲学从古希腊到现在的一个中心主题是人类自由意志。在量子力学之前的牛顿定律和大部分物理定律都有确定性的方程。几乎所有西方哲学家和大部分科学家都将此解读为宇宙是确定性的。康普顿（The Freedom of Man，1935；The Human Meaning of Science，1940）和波普尔（“云和钟，”1972）都有力地论证了量子力学终结了这种世界观。虽然如此，这个思想仍然被认为有哲学和/或科学基础。

[7] Steven Pinker, How the Mind Works（1997）很好地阐释了大脑的功能，这本书极具可读性。

[8] Donald T. Campbell，“进化认识论”，收录在The Philosophy of Karl Popper, P.A.Schlipp ed.，pp.412—463（1974），第12章。

[9] 杰拉尔德·埃德尔曼因抗体结构的研究获得了1972年诺贝尔生理学或医学奖，随后35年里致力于研究大脑工作原理。他写了几本书介绍神经达尔文主义和意识。A Universe of Consciousness, How Matter Becomes Imagination（2000）是很好的入门读物。Neural Darwinism：A Theory of Neuronal Group Selection（1987）详细介绍了他的思想。

[10] 贝叶斯大脑的思想综合了20年来的几条思路。近年来贡献最多的是伦敦大学学院的Karl Friston。我不知道有没有好的大众读物，如果想进一步了解的话可以参考3篇综述文章：“The Bayesian Brain：The Role of Uncertainty in Neural Coding and Computation.”David C.Knill&Alexandre Pouget, TRENDS in Neuroscience 27：712—719（2004）；“The Free Energy Principle.”Karl Friston, Nature Reviews Neuroscience 2：127—138（2010）；以及“Hierarchical Bayesian Inference in the Visual Cortex.”Tai Sing Lee&David Mumford.J.Opt.Soc.Am A.，20：1434—1448（2003）。

[11] Karl Friston，“A Theory of Cortical Responses，”Philosophical

Transactions of the Royal Society 360：815—836（2005）以及注释10提到的“The Free Energy Principle”（2010）。

[12] 要了解埃德·索贝博士的工作，参见网站www.invention-center.com。

第11章

[1] 在《自私的基因》（1976）中，道金斯讨论了群体遗传学的基本原理，并提出了拟子的概念。

[2] 埃里克·拜因霍克在《财富的来源》（The Origin of Wealth，2006）一书中解释了扩增和演绎修补（deductive tinkering）的作用。这本书很好的阐释了经济是进化的复杂适应系统的思想。

[3] 事实上，大部分实践科学家对这个主题并不是很关心。科学家们几乎从不根据科学哲学来做研究，而是向成功的科学家学习。《科学发现的逻辑》的作者卡尔·波普尔和《科学革命的结构》的作者托马斯·库恩都不是实践科学家；波普尔是哲学家，而库恩则是历史学家。

[4] 有大量文献将科学视为进化过程。重要的作者包括卡尔·波普尔、唐纳德·坎贝尔、斯蒂芬·图尔明和大卫·赫尔。我推荐波普尔的《科学发现的逻辑》和大卫·赫尔的Science as a Process, an Evolutionary Account of the Social and Conceptual Development of Science（1988）。

[5] 埃里克·拜因霍克，《财富的来源》，2006。

[6] 同上。丹尼特，《达尔文的危险思想》，

[7] 1995。埃里克·拜因霍克，《财富的来源》，

[8] 2006.

[9] David Sloan Wilson, Darwin's Cathedral：Evolution, Religion, and the Nature of Society（2002）.

第12章

[1] Christoph Adami, Charles Ofria，&Travis C. Collier.“Evolution of Biological Complexity，”in Proceedings of the National Academy of Sciences（2000）。这篇学术论文将复杂性定义等同为信息论的负熵，并证明这个量在特定的进化系统中增加。我反对这样使用复杂性一词，因为这样刻画的是特定系统与最优的距离，与通常的复杂性意义很不一样。

[2] 这里的结构和功能复杂性的区别引自：Francis Heylighen, The Growth of Structural and Functional Complexity During Evolution，收录在The Evolution of Complexity（1999）。海利戈恩对复杂性及其各种定义方式多有论述。

[3] 同上。

[4] 关于深度最重要的学术论文：Charles Bennett, Logical Depth and Physical Complexity，收录在The Universal Turing Machine：A Half Century Survey, Oxford University Press，1988。

[5] 图12.1中的数据来自：J.J.Sepkoski, A Kinetic Model of Phanerozoic Taxonomic Diversity.Ⅲ.Post-Paleozoic Families and Mass Extinctions, Paleobiology（1984）。这些数据被广泛引用、讨论和分析。陆生动物和维管植物也有类似的数据报道（P.W.Signor, American Zoologist\[1994\]）。虽然被多次修正，但多细胞生物的多样性在长时期里增长这个基本结论已经得到了化石记录的证明。

[6] 斯蒂芬·古尔德喜欢强调地球上的生命完全依赖于细菌，而不怎么依赖于哺乳动物的存在。在他的书中经常有这样的观点。

[7] Steven J. Gould, Full House：The Spread of Excellence from Plato to Darwin（1996）.

[8] Daniel McShea, Complexity and Evolution：What Everybody Knows, Biology and Philosophy（1991）。麦克谢伊可能比其他同时代的生物学家更深入地研究了进化与生物复杂性之间的某些关联。

[9] 图12.3最早出现在：T.R.Gregory, Macroevolution, Hierarchy Theory, and the C-value Enigma, Paleobiology（2004）。动物基因组大小

数据库有最新数据，http：//www.genomesize.com/index.php。

[10] 图12.4经许可重绘，引自：James W.Valentine, Allen G.Collins，&C.Porter Meyer, Morphological complexity increase in metazoans, Paleobiology 20：pp.131—142，1994。更近的数据表明曲线前部的弯曲是数据不完整所致。目前对多孔动物门起源时间的估计大约为7亿年前，刺胞动物门约为6.8亿年前。（有兴趣的读者可以参考：The Cambrian Conundrum：Early Divergence and Later Ecology Success in the Early History of Animals, D.H.Erwin, et al.，（2011），Science 334，pp.1091—1097。）

[11] Stuart Kauffman, The Origins of Order（1993），以及更浅显的At Home in the Universe（1995，《宇宙为家》，湖南科学技术出版社）。

[12] McShea，1991.

[13] 丹尼尔·阿什洛克，个人信件。5个发表的例子参见：Dan Ashlock&John E.Mayfield, Acquisition of General Adaptive Features by Evolution, Evolutionary ProgrammingⅦ，V.W.Porto, N.Sarava-nan, D.Waagan，&A.E.Eiben eds.（1998）；Daniel Ashlock, Elizabeth Blanken-ship，&Jonathan Gandrud, A Note on General Adaptation in Populations of Painting Robots, Proceedings of the 2003 Congress on Evolutionary Computation（2003）；Daniel Ashlock&Brad Powers, The Effect of Tag Recognition on Non-Local Adaptation, Proceedings of the 2004 Congress on Evolutionary Computation, Vol.2（2004）；David Doty, Non-Local Evolutionary Adaptation in Gridplants, Proceedings of the 2004 Congress on Evolutionary Computation, vol.2（2004）；Daniel Ashlock&Adam Sherk, Non-Local Adaptation of Artificial Predators and Prey, Proceedings of the 2005 Congress on Evolutionary Computation（2005）。

[14] W. Ross Ashby, An Introduction to Cybernetics（1956）。参

[15] 见注释4。

[16] Leigh Van Valen, A New Evolutionary Law, Evolutionary

Theory（1973）。红皇后假说认为生物不断与其他进化的生物竞争，并且每个类群（种群、类属，等等）的环境因为竞争者的进化也在不断变化。在这样的环境中，生存的唯一途径就是不断进化。这个假说被用来解释在大量的类群中，灭绝的可能性一直存在，并且与群体已经存在的时间无关。一旦类群无法通过进化适应其竞争者的变化就会灭绝。

[17] Lathrop, Computing and Evolving Variants of Computa-tional Depth，1997，博士论文，以及Compression depth and genetic programs, Proceedings of the Second Annual Conference on Genetic Programming, pp.370—378.Morgan Kaufmann, San Francisco，1997。普通读者可能会觉得这些文章难以理解，但能大致了解一下理论计算机科学的抽象性。

[18] Bennett（1988）将逻辑深度比喻为一个对象的内在“价值”，并且将价值直接联系到“证明一个对象的计算历史的量”。Juedes（1994）等人所说的“计算有用性”，Lutz（个人信件）说的“冗余的组织”，以及Lathrop（1997）引用的Bennett和Juedes的说法，认为深度表现了“繁杂结构”，“x的组织中存储的计算工作”，和“通过计算嵌入在字符串中的组织”。因此，我们看到本内特所说的价值是基于这样的观念，即输入的信息可以被计算过程组织成对将来的计算越来越有用的输出结构。

第13章

[1] 有许多书论述生命的起源。最适合入门但又很老的一本书是Robert Shapiro的Origins, A Skeptic’s Guide to the Creation of Life on Earth（1987）。还有两本较新的书是William Schopf编辑的Life’s Origin（2002）和Andrew H.Knoll的Life on a Young Planet, the First Three Billion Years of Evolution on Earth（2003）。2009年公布了一项重大突破，John Southerland在曼彻斯特大学的团队发现了一条化学合成嘧啶核苷酸的可行路径，这种物质是RNA的前体（Powner, M.，Gerland, B.，&Sutherland, J.，Nature[May 2009]）。此前核苷酸的合成被认为是生命的前生物化学起源最大的难题。目前依然不知道在同样的原汤中产生氨基酸、核苷酸和糖的条件，但Powner等人发现的途径在这个方向上推进了一大步。

[2] 两种成分RNA系统的持续复制在2009年被发现。在这个系统中，两种预先合成的具有RNA连接酶活性的RNA与4种特别设计的RNA

聚合物结合时会产生至少两种连接酶。如果加入这4种前体聚合物，会同时产生越来越多的连接酶。只要有前体分子，这个系统似乎能一直复制下去（T.Lincoln&G.Joyce, Science[February 27，2009]）。

[3] Stuart Kauffman, Reinventing the Sacred（2008）.

[4] Bernd M. Rode, Daniel Fitz，&Thomas Jakschitz, The First Steps of Chemical Evolution Towards the Origin of Life, Chemistry&Biodiversity（2007）。这篇综述文章很好地介绍了各种化学起源理论，但没有包括2009年发表的一些突破（参见注释1和2的文献）。

[5] Stuart Kauffman的两本书阐释了这个思想：Reinventing the Sacred（2008）和Origins of Order（1993）。专业性的分析参见E.Mossel&M.Steel, Random Biochemical Networks：The Probability of Self Sustaining Autocatalysis, Journal of Theoretical Biology（2004）。

[6] Harold J. Morowitz, Vijaysarathy Srinivasan, Shelley Copley，&Eric Smith, The Simplest Enzyme Revisited, the Chicken and Egg Argument Solved, Complexity（2005）.

[7] 在化学中摩尔的定义是物质所含成分（例如原子、分子、离子、电子）的数量等于12克碳12的原子数量的物质的量。1摩尔纯物质含有6.0221415×1023个原子或分子。1摩尔所含的质量的克数等于物质的分子/原子量。因此，可以通过称重并与分子/原子量比较来测量纯物质的摩尔量。这个解释来自维基百科，在任何一本化学课本中都有类似的定义。

[8] Matt Inlay, Evolving Immunity, A Response to Chapter 6 of Darwin's Black Box。这篇文章的网址位于http：//www.talkdesign.org/faqs/Evolving_Immunity.html。

第14章

[1] 维基百科对摩尔定律有很好的总结，给出了参考文献和数据表。

[2] Ray Kurzweil, The Singularity Is Near（2005）。这本书的许多图可以在书的网站上看到，singularity.com。

[3] A. M.Fennimore, T.D.Yuzvinsky, Wei-Qiang Han, M.S.Fuhrer, J.Cumings，&A.Zettl, Rotational Actuators Based on Carbon Nanotubes, Nature（2003）。网上有这种马达运行的视频：http：//www.livescience.com/technology/050412_smallest_motor.html。

[4] D. M.Eigler&E.K.Schweizer, Positioning Single Atoms with a Scanning Tunneling Microscope, Nature（1990）。在IBM STM图库中还有其他原子的图像：http：//www.almaden.ibm.com/vis/stm/atomo.html。

[5] 美国国立卫生研究院对人工耳蜗的移植提供了简单介绍：http：//www.nidcd.nih.gov/health/hearing/coch.asp。随着这种设备的商业化，一些供应商在网上也提供了教育材料。

[6] 实际上这是个很怪的图，随着时间流逝，数字4的点会上移接近期望线。这个点落在所画的线上需要数百年时间。要进行推测有一个基本问题是，只有很少的数据点时无法确切得知到底在哪里划线。

名词解释

ANN

　　参见人工神经网络。

B细胞（B-cell）

　　能产生抗体的细胞。

CAS

　　参见复杂适应系统。

DNA

　　脱氧核糖核酸的缩写，基因的分子结构。所有活细胞都有（或曾有）编码细胞生长和分裂的必要信息的DNA。DNA分子是由脱氧核苷酸组成的长链。一些DNA分子由数百万个脱氧核苷酸组成。

DNA聚合酶（DNA polymerase）

　　以现有DNA为模板，将脱氧核苷酸连接到一起合成DNA分子的酶。

DNA序列（DNA sequence）

　　构成DNA分子的脱氧核苷酸的特定序列。

II型事物

　　形成过程需要指令的事物。

IPCS

选择性迭代概率计算，复杂引擎的一种更正式的说法。

Ⅰ型事物

形成过程不需要指令的事物。

MHC

参见主要组织相容性。

mRNA

参见信使RNA。

RNA

核糖核酸的缩写。核苷酸的多聚体，结构非常类似于DNA，但在细胞中有不同的功能。类似于DNA, RNA也编码信息。

RNA剪接

细胞中许多RNA分子最初是合成为更大的前身分子。改变这些前身分子的方式之一是去掉至少一个内含子。去掉内含子并重连RNA分子就称为剪接。

RNA聚合酶（RNA polymerase）

将核苷酸连到一起合成RNA的酶。

T_4

一种噬菌体。

T细胞

一种能产生T细胞受体的免疫系统细胞类型。一些T细胞能检测和摧毁体内受病毒感染的细胞；还有一些则调控不同免疫系统细胞的行为。

T细胞受体（T-cell receptor, TCR）

T细胞表面产生的蛋白质。T细胞受体结合细胞膜上MHC蛋白质“表达”的外来多肽（氨基酸链）抗原。参见MHC。

ZIP文件（ZIP file）

使用ZIP压缩算法压缩的文件。ZIP文件比源文件小。源文件可以通过将ZIP文件解压恢复。

α-粒子（Alphaparticle）

由两个质子和两个中子组成的原子。

氨基酸（Amino acids）

蛋白质分子的化学构成。氨基酸可以单独存在也可以连接成链。

氨基酸侧链（Aminoacid side chain）

氨基酸的中心结构由两个碳原子和一个氮原子结合在一起组成。端部的碳是“酸”基的一部分，中间的碳与氢原子以及另一个化学结构结合，氮原子是氨基的一部分。与中间的碳连接的化学结构称为侧链。

白细胞（Leukocyte）

血液或组织中的免疫细胞，也称为白血球。

贝纳德流（Bénard cells）

将薄层液体从底部或顶部均匀加热形成的规则循环对流模式。

贝叶斯（Bayesian）

由托马斯·贝叶斯开创的统计学分支。这种数学理论融入了存在先验概率分布的思想（在收集任何数据之前概率就已经存在）。

贝叶斯最优（Bayes optimal）

从贝叶斯统计的角度来看最优的统计结果。

编码（Code）

将一种信息形式转化成另一种形式的规则。例如，摩斯码将点线形式和字母相互转化。基因编码将DNA中的核苷酸序列转化成蛋白质中的氨基酸序列。

表型（Phenotype）

生物学概念，泛指生物的生理特征。

布尔（Boolean）

从乔治·布尔（1854）的著作中衍生出来的数学或逻辑概念。布尔函数可以取两个值：真或假（1或0）。

程序（Program）

计算机可以执行以完成特定任务的指令。

催化

由于某种物质的参与导致化学反应速率的变化称为催化。化学反应不会消耗反应的催化剂。催化剂通常能让反应在更低的温度下进行。在生物学中，酶是对特定反应具有高度特异性的蛋白质催化剂。

单体（Monomer）

小的化合物单元，是名为多聚体的大型分子的组成模块。多聚体由多个单体组成。参见多聚体。

蛋白质（Protein）

由一条或多条氨基酸单体链组成的生物多聚体。蛋白质负责细胞中的大部分活动，包括结构的创建和化学反应的催化。

蛋白质合成（Protein synthesis）

蛋白质在细胞中的组装过程。蛋白质中的氨基酸序列完全由相应的信使RNA的核苷酸序列决定。

蛋白质组（Proteome）

细胞中蛋白质相互作用的网络。

点突变（Point mutation）

涉及一个比特位或一个核苷酸的自发突变。

迭代（Iteration）

反复进行的过程，每次循环的结果作为下一次循环的起点。

定律（Law）

通常为简单的数学关系，一般很成熟，但不一定。

动量（Momentum）

运动质量沿相同方向相同速度持续运动的趋势。动量的值等于质量与速度的乘积。

动作电位（Actionpotential）

特殊的离子流模式，伴随着沿神经元（神经细胞）传递的电压变化。动作电位是神经元与其他神经元和肌肉细胞交流的主要方式。

对流环（Convectioncell）

液体或气体底部加热顶部冷却时出现的协同运动。暖液（气）流上升，冷液（气）流下降，形成规则模式。贝纳德流就是一种对流环。

多聚体（Polymer）

由大量相似或相同的“单体”连接到一起形成。就好像珍珠项链。

多肽（Polypeptide）

氨基酸链，在本书中这个词可以与蛋白质互换。

多肽自发折叠（Spontaneous folding of apolypeptide）

氨基酸链自发折回自身形成紧凑的结构。链的3维结构取决于氨基酸以及周围液体的相互作用。通过这种方式，所有氨基酸序列都有特有的紧凑形状。

多糖（Polysaccharides）

大分子多聚体，单体是单糖（糖）。

反向传播（Back propagation）

一种训练人工神经网络的方法。这种方法需要知道所期望的网络输出数值，并且将节点加权输入转化为输出的数学函数（激活函数）是可微的（在微积分中，可微函数在定义域的任意点都存在导数）。

非平衡结构（Nonequilibrium structure）

不处于平衡态的系统中出现的结构。这类结构通常需要有通过系统的持续能量流。没有能量流，结构就会耗散。

分子（Molecule）

由至少两个原子通过化学键结合组成的结构。分子通常具有与其原子成分完全不同的性质。

分子拼砌（Moleculartiling）

分子砖组装到一起形成结构的过程。

分子运动（Molecularmotion）

与热运动相同。分子的自发运动，表现为温度。

分子砖（Molecular tiles）

人造的化学结构，通常由DNA构成，置于溶液中时能通过自发组装物理结构执行计算。

封闭系统（Isolatedsystem）

在热力学中，如果没有能量（或物质）能进入或离开系统，就说系统是封闭的。

符号序列（Sequences of symbols）

符号组成的一维序列。

负熵（Negentropy）

系统的熵与系统的最大（平衡态）熵之间的差值。

复杂适应系统（Complex adaptive system）

各部分之间互动，并且对过去的互动具有（一定的）记忆的系统。在这类系统中，未来的互动部分取决于过去的互动。

复杂性（Complexity）

对事物的复杂程度进行量化的未严格定义的概念。事物的复杂性通常被认为是系统或对象的组分数量以及组分之间的各种互动的某种函数。

复杂引擎（Engine of complexity）

所有（广义达尔文意义上的）进化过程都具有的计算循环。框图见图5.2。这个循环一旦实现，将能组装复杂的指令。

概率计算（Probabilistic computation）

至少包括一个随机步骤的计算。即使每次运行的输入相同，概率计算每次运行也会产生不同的结果。

构造引擎（Engine of construction）

类似于复杂引擎的循环，只是在选择之前没有复制。这个循环能产生惊人的事物，但不像复杂引擎那样具有创造性。

还原性条件（Reducing conditions）

还原因子占主导的状态（参见氧化/还原）。在生物圈中最常见的氧化因子是氧分子，最常见的还原因子则是没有与氧结合的氢。还原性大气包含氢分子、氨（氮与氢结合）、甲烷（碳与氢结合）、和少量氧分子。

海森伯测不准原（Heisenberg uncertainty principle）

原子尺度的粒子特定的物理属性对无法同时确定。一个例子是对粒子在空间中的位置测量越精确，对动量的测量就越不精确，反过来也是一样。

寒武纪（Cambrian period）

5.42　亿年前到4.88亿年前的一段地质时期。寒武纪地层是具有明显动物化石的最古老的地层。在比寒武纪更古老的地层中，动物化石很小而且罕见，很难与现代动物学分类关联。

寒武纪海洋（Cambrian sea）

寒武纪地质时期的海洋。

核苷酸（Nucleotide）

RNA的单体化合物，DNA的单体通常也这么称呼。DNA的单体的正式名称是脱氧核苷酸。参见单体。

核苷酸链（Nucleotide chain）

核苷酸组成的链，其中一个核苷酸的糖基与下一个核苷酸的磷酸基通过化学（共价）键相连。

核糖体（Ribosomes）

细胞中包含的分子结构，是细胞中合成蛋白质的场所。

红皇后（Red Queen）

在协同进化系统中，每个物种都必须不断进化，以维持相对于系统中其他物种的适应性。这个词源自刘易斯的《爱丽丝镜中奇遇记》中的红皇后赛跑。皇后在爱丽丝身边边跑边说，“你看，你必须全力奔跑，才能保持不动。如果你想去别的地方，奔跑的速度就必须至少再快一倍。”

宏观态（Macrostate）

物理系统在比原子大的尺度上的状态。系统的宏观态由压力、温度、和外观等属性刻画。

化学不稳定（chemical Labile）

不稳定的化合物很容易分解。

化学性（Chemical）

与原子或分子相互作用有关的现象或条件。

混沌理论（Chaos theory）

分析对初始条件具有敏感依赖性的动力系统的行为的数学分支。

基因（Gene）

DNA的功能单元。大部分时候这个词指的是编码形成某种蛋白质所必需的信息的DNA段。

基因表达（Gene expression）

基于基因序列生成某种蛋白质。

基因网络（Genetic network）

在细胞中一些基因生成的蛋白质调控其他基因蛋白质的生成，这样形成的调控网络称为基因网络（基因通过生成的蛋白质调控其他基因）。

基因型（Genotype）

生物的基因组构成的类型。

基因重组（Geneticrecom bination）

DNA分子断开和重连的细胞过程，新的DNA分子由两个或更多原来的DNA分子的段组成。

基因组（Genome）

一个生物的所有DNA序列。所有生物（除非是同卵双胞胎或克隆）都有独一无二的基因组。

脊索动物（Chordates）

脊索动物是一种包括脊椎动物在内的大的动物门类，也包括尾索动物和头索动物。脊索动物都有被称为脊索的发育结构，是脊椎动物背脊骨的前身。

计算（Computation）

输入通过设备产生相应输出的过程。计算通常有多个步骤。

计算链条（Computational chains）

前后相连的计算，一个计算的输出作为下一个计算的输入。

钾通道（Potassiumchannel）

让钾离子而不让其他离子通过生物膜的蛋白质结构。

简单计算（Computationally simple）

很容易从短输入计算出来。这里的容易指的是计算只需少量步骤和计算资源。

简单输入（Simple input）

一般表示短输入。如果输入不是序列，则表示输入规模较小。

胶束（Micelle）

由一大类结构类似洗涤剂的分子形成的结构。这类分子的一端亲水，另一端疏水。当置于水中时倾向于形成拉长的球体（胶束），亲水端朝外，疏水端朝内，将水排斥在胶束外面。

节点（Node）

网络的汇合点。节点可以是与其他节点连接或互动的任何东西。

结构（Structure）

同有序一样，与随机相对。在结构中各部分有固定（或者至少不随机）的关系。

进化（Evolution）

这个词有多重意义。最简单的是进步的变化。这本书极少使用这个意义；它定义得太模糊。另一个是达尔文式变化，意指生物谱系的遗传变化。第三个意义，也是这本书中最常用的，是与复杂引擎有关的变化。这是广义的不限于生命系统的达尔文式变化。

进化发育生物学（EvoDevo）

试图基于分子调控机制推演或重构生物结构的进化史的生物学分支。

进化计算（Evolutionary computation）

利用了复杂引擎的计算。

进化认识论（Evolut ionary epistemology）

一种哲学流派，认为知识以类似自然选择的过程进化。参见认识论。

进化适应（Evolved fit）

适应的出现有多种途径。如果是通过复杂引擎的方式，就称为进化适应。

进化算法（Evolutionary algorithm）

包括“遗传算法”在内的一类计算机算法，其中融合了复杂引擎。

静息电位（Restingpotential）

未受刺激的情况下跨细胞膜的电位。

局部最小/大（Localminimum/maximum）

在具有峰谷的地形中，有一个山谷是最低点，一个山峰是最高点。局部最小是比最低点高但周围的地势都相对更低的山谷。

巨噬细胞（Ma—crophage）

白细胞的一种，类似阿米巴虫，在体内巡逻，吞噬碎屑和细菌。

抗体（Antibody）

适应性免疫系统产生的特殊蛋白质。当体内出现外来分子时，抗体能高度特异性地与其结合。抗体由B细胞产生，在血液和淋巴中循环。有些抗体也可以通过分泌产生。

抗原（Antigen）

免疫系统响应的分子（通常是蛋白质或多糖）。通常这种响应是抗

体或T细胞受体与抗原的结合。

柯尔莫哥洛夫复杂性（Kolmogorov complexity）

参见算法信息。

可数性（Countability）

（数学中的）可数集指的是元素个数不多于自然数的集合。不可数集指的是元素个数多于与自然数一一对应的集合。可数集具有可数性。

库仑定律（Coulomb’slaw）

两个点电荷之间的斥力正比于电荷大小的乘积，反比于两者之间距离的平方。

跨膜通道（Membrane channel）

嵌在细胞膜上选择性允许小分子或离子通过膜的蛋白质。

扩增（Amplification）

埃里克·拜因霍克尔用这个词描述在社会的进化中，商业计划等富含信息的产物的传播和增殖。

离散（Discrete）

不连续；颗粒化；元素具有分离的值。在离散系统中，元素相互区隔。

离子（Ion）

具有（正的或负的）净电荷的原子或分子。

离子泵（Ion pump）

嵌在细胞膜上能转移特定离子通过膜的蛋白质。泵需要化学能来源以执行动作。

离子通道（Ion channel）

嵌在细胞膜上能选择性让特定离子通过膜的蛋白质。

理论（Theory）

对现象的解释。理论可以是猜测性的，也可以是严格确证的。

理想气体定律（Universal gas law）

气体的压力、体积和温度有如下关系：PV=nRT。其中P是压力，V是体积，T是温度，R是决定单位的常数，n是气体的量。

量子力学（Quantum mechanics）

描述和解释极小尺度的物理现象的数学理论。

临界状态（Critical state）

小的扰动能引发大的变化的系统状态。

磷脂（Phospholipid）

生物膜的主要成分。

磷脂双层（Phospholipid bilayer）

磷脂形成的双层结构。磷脂双层是生物膜的基本结构。

逻辑门（Logic gate）

实现布尔函数的物理设备。参见布尔函数。

慢增长（Slow growth）

深度的特性。深度对象要从简单（短）输入产生，只能通过深度计算。

酶（Enzyme）

对特定反应具有高度特异性的蛋白质催化剂。酶也调控它们催化的化学反应。自然界的酶都由活细胞产生。参见催化。

密码子（Codon）

mRNA中3个相邻核苷酸的组合。参见信使RNA。4种核苷酸3个一组总共有64种排列（AAA、AAG、ATC，等等）。每种排列称为一个密码子。其中61种对应蛋白质中特定的氨基酸，还有3种是“终止”密码子。

幂律（Power law）

一种数学关系，描述事件的频率正比于事件规模的某次幂。当现象遵循幂律，小事件比大事件要常见得多。

免费的结构（Structure for free）

形成结构无需遵循指令。

免疫（Immunity）

生物依靠免疫系统建立起稳固的分子和细胞防疫，保护生物免受特定抗原的感染。免疫包括天生免疫和适应性免疫。

模板（Template）

用来形成互补结构的物理结构。在DNA复制过程中，DNA双螺旋的一条链作为新链的模板。这个词与元胞自动机中的种子行具有相同的意义。

摩尔定律（Moore’s law）

新发布的微处理器的晶体管数量每两年增长一倍的经验法则。这个法则在40多年里都一直有效。

目的性信息（Purposeful information）

具有意向性的结果或产物的特定信息组织。目的性信息是“为了”某事物。

钠离子通道（Sodium channel）

允许钠离子通过生物膜的蛋白质结构。

囊泡（Vesicle）

在生物学中，由生物膜形成的封闭球状结构称为囊泡。囊泡中含水。人工磷脂双层也可以形成囊泡。

能量（Energy）

能量和熵是经典热力学的两个基本属性。能量与质量以及变化的势有关。第一热力学定律说的是能量既不能创生也不能消灭，但可以改变形式。

能量守恒（Conservation of energy）

热力学第一定律：能量既不能创生也不能消灭（但是可以转换形式）。参见能量。

拟子（Meme）

在文化中，通过人际传播的思想、行为或风尚。这个词是理查德·道金斯在《自私的基因》中提出，用于讨论文化的进化。拟子类似于遗传学中的基因，但没有同等的结构或精确定义。

牛顿万有引力定律（Newton’s Law of gravity）

物质相互吸引，两个物质（点）的引力正比于质量的乘积，反比于两者距离的平方。

配子体（Gametophyte）

能产生精子或卵细胞的植物或植物组织。植物表现单倍体（一份染色体）繁殖和二倍体（两份染色体）繁殖交替的生命循环。单倍体繁殖

的植物称为配子体；二倍体繁殖的植物称为种子植物。几乎所有常见的植物都是种子植物，对于种子植物，配子体是花粉（雄性）或花中的特定组织（雌性）。

平衡态结构（Equilibrium structure）

系统处于平衡态时所呈现的结构。这种结构不会变化，但一些会表现出重复运动，就像钟摆或围绕太阳运转的行星。

启动子序列（Promoter sequences）

细胞中蛋白质转录因子结合的DNA段。转录因子的结合会改变RNA聚合酶启动相邻DNA段的RNA副本的合成。参见转录因子。

前身分子（Precursor molecules）

分子都是由化学反应产生。在化学反应中前身分子反应生成产物分子。前身分子也称为反应物。在生物细胞中，反应物本身又是其他反应物的化学反应产物。从而在最后的分子产生之前有很长的反应链条。这个链条上的任何反应物都可以认为是最后的产物的前身分子。

氢键（Hydrogen bond）

一个分子中与氧或氮原子（共价）结合的氢原子同另一个分子中与氧或氮原子结合的氢原子之间相互吸引。氢键将不同的分子（或大分子的不同部分）连接到一起，强度约为共价键的1/10。

确定性（Deterministic）

完全由状态决定，没有选择，没有不确定性。

确定性混沌

这个词指的是可以用确定性方程描述的不可预测的行为。对于这

（Deterministic chaos）

种情形，如果要预测很远的未来，对初始条件的测量需要有无穷精

度。参见混沌理论。

确定性计算（Deterministiccomputation）

相同输入总是产生相同输出的计算。

群体遗传学（Population genetics）

遗传学的分支，研究混杂繁殖生物群体中基因频率的变化。群体遗传学从数学上描述生物进化的原理。

热力学（Thermodyna-mics）

研究能量和做功的物理学分支。

热力学第二定律（Second law of thermodynamics）

封闭系统的熵要么增加要么保持不变。另一种说法是热量总是从热往冷传递。

热力学第一定律

能量既不能创生也不能消灭，但能改变形式。参见能量。

热力学平衡（Equilibrium in thermodynamics）

如果外界没有输入，系统宏观上就不会有变化的状态。在平衡态所有力和势都处于平衡。

热运动（Thermal motion）

微小结构的自发运动，温度的物理表现。

人工神经网络（Artificial neural network）

一种计算模型，由几层简单的处理单元（节点）组成，每个节点从前一层的节点获得输入。第一层接受输入数据，每个节点将输出传送到后一层的节点。节点输出是对节点的所有输入进行加总或其他简单变

换。典型的ANN有3层。

人工神经网络权重（Weights of an ANN）

人工神经网络的节点有多个输入，但输入的权重不一定相同。一些输入相对来说影响更大。权重是与输入值相乘的数值，决定输入有多重要。权重为0意味着输入被完全忽视。参见人工神经网络。

人类中心论（Anthropocentrism）

认为人类最重要，或者以人类为中心的看问题的立场。

认识论（Epistemology）

关注知识的本质和范畴的哲学分支。

冗余（Redundancy）

系统中的重复。在计算机科学中，具有冗余的数字对象可以被压缩而不会损失信息。

塞平斯基三角形（Sierpinski gasket）

由各种大小三角形组成的图样，通过反复应用异或布尔函数生成。参见异或。

三元超级网（Tripartite supernetwork）

参见超级网络。

山峰函数（Hill function）

画成3维很像对称的山峰的数学函数。

熵（Entropy）

熵和能量是经典热力学的两个基本属性。当能量一定，熵度量的是无做功能力的程度。在统计力学中，熵度量的是系统的随机或无序的程

度。当系统处于平衡态，系统的熵处于最高值，系统无法做功。参见香农熵。

深度（Depth）

在这本书中意指几个数学度量，包括“逻辑深度”（丹尼特）和“计算深度”（卢茨）。通俗的说，深度度量的是从简单输入计算出数字对象的困难程度。

神经元（Neuron）

神经细胞。身体中组成神经和神经节的细胞，并且负责身体中的长距离通信。

神经元群选择（Neuronal group selection）

杰拉德·埃德尔曼提出的大脑形成和功能的理论。在这个理论中，大脑的总体生理结构由遗传决定，但神经元连接的细节由生长和发育过程中的细胞选择决定。在这个过程中，神经元群自组织成功能模块。在大脑初步形成后，突触选择继续作用于功能模块以及单个神经元。

神经元通信模式（Neuronal communication patterns）

生物神经元相互之间的通信网络。

生物钟（Circadian clock）

以大约24小时为周期的生物化学反馈环。大部分细胞都至少具有一个生物钟。

适应性地形（Adaptive landscape）

以3维方式展现某个问题的许多可能答案的值的方式。画出来的适应性地形很像卡通中的山地。

适应性免疫（Adaptive immunity）

体内出现感染后被激活的免疫系统部分。激活的免疫系统针对身体

遇到的特定疾病产生出针对性的分子机制。

噬菌体（Bacteriophage）

一种感染细菌细胞的病毒。

收敛式进化（Convergent evolution）

最初很不一样的生物进化出很类似的特征。典型例子包括世界不同区域的沙漠植物都进化出了能进行光合作用的茎干，叶子则变成了不进行光合作用的刺。

双螺旋（Double helix）

细胞中DNA的分子形状。双螺旋由两条DNA分子单链相互环绕而成，中间通过多个氢键相连。两条DNA链的核苷酸序列互补，如果一边是“A”另一边则是“T”，一边是“G”另一边则是“C”。

算法（Algorithm）

执行某个过程或解决某个问题的一组指令。算法通常指的是能让计算机执行特定动作的一系列命令。

算法随机

不可压缩，无论何种计算都无法减少不可压缩对象的大小。

算法信息

在计算机科学中很有用的一种信息度量。比特位串的算法信息内容可以通俗地认为是输入通用计算机后能生成该位串的最短可能的位串的长度。

随机变异（Random varation）

生物学中指的是没有预设效果的DNA变异。

随机的（Stochastic）

概率性的，部分或完全随机；具有随机的要素。

随机数发生器（Random number generator）

通常指的是输出随机数的算法。如果计算机是确定性的（几乎所有计算机都是），则输出的是伪随机数。

随机性（Randomness）

当前或之前状态的知识不能为将来的事情提供信息。

随机序列（Random sequence）

顺序不可预测的符号列。

天生免疫（Innate immunity）

在生物接触特定的致病生物之前就存的防疫感染的机制。由于有天生免疫，大部分可能的感染都没有发生。

停机问题（Halting problem）

在计算机科学中有时候无法知道正在进行的计算是否最终会产生输出还是会一直进行下去永不产生输出。

通用计算机（Universal computer）

能编程模拟其他任何计算机的计算机。个人电脑就是通用计算机。

突变（Mutation）

生物DNA或遗传算法编码的自发或随机变化。

突变率（Mutation rate）

突变发生的频率。

图灵机（Turing machine）

阿兰·图灵于1936提出的一种假想的机器，用于证明计算机科学中的一些定理。图灵机根据少量简单的规则操作条带上的符号。它是现代计算机的思想先驱。

脱氧核苷酸（Deoxynucleotide）

DNA的化学单元。脱氧核苷酸通常也简称核苷酸。

网络（Network）

由相互连接的节点组成的结构。节点可以是任何东西。

网络的稳定状态（Stable states of a network）

许多网络都具有动态行为。当动态网络进入不变的状态，就称为稳定状态。外界输入经常会改变或破坏稳定状态，导致向新状态的转移。

微观态（Microstate）

系统在原子或分子尺度上的所有瞬时位置和动量。

物质（Matter）

宇宙中具有静质量和体积的“东西”。通常这意味着其由原子组成。物质具有质量，质量又通过爱因斯坦质能方程$E=mc^2$ 与能量相关联（其中E是能量，m是质量，c是光在真空中的速度）。

系统（System）

任何具有边界的事物。

细胞（Cell）

生物的基本组织单元。细胞被生物膜包裹，DNA决定其中表达的蛋白质。

细胞的超级网络（Supernetwork in cells）

细胞具有生物化学反应网络、基因调控网络以及蛋白质相互作用网络（蛋白质组）。这三个网络组成了界定生物细胞的超级网络。

细胞质（Cytoplasm）

细胞中位于细胞膜和细胞核（如果细胞有核）之间的成分。

香农度量（Shannon measure）

香农信息。

香农熵（Shannon entropy）

香农信息。

香农信息（Shannon information）

香农基于各种可能性的概率提出的信息的数学定义；与统计力学的熵的定义密切相关。

协同进化（Coevoution）

至少2种物种一起进化，物种进化以应对其他物种的进化。典型的例子包括，羚羊进化得速度越来越快，以逃脱豹子的追捕，而豹子为了追逐羚羊也进化得速度越来越快；花进化出红色管状花以吸引蜂鸟，蜂鸟则进化出长喙从管状花中吸取花蜜。

新科学系统（NKS systems）

NKS是“新科学”的缩写。这个词是由斯蒂芬·沃尔弗拉姆提出，用来描述用简单规则和输入产生出明显的复杂性的逻辑系统。

信使RNA（Messenger RNA, mRNA）

RNA的一种形式，将DNA序列复制为核糖体，RNA的核苷酸序列决定组成多肽的氨基酸序列。

信息（Information）

一个简单的定义是“符号模式传递的意义”。更严格的数学定义是香农信息和算法信息。参见香农信息和算法信息。

信息体（Information body）

为了传递意义以某种方式组织的符号集。

形态（Morphological）

形状或形式。

选择性剪接（Alternative splicing）

细胞中的一种修改RNA的方式。在剪接过程中RNA分子在两点被截断，中间部分被移除。当同一个RNA发生多处剪切时，剪切点可以以不同方式连接，导致RNA不同的中间区域被移除。剪接最终会产生怎样的RNA取决于剪切点被怎样组合到一起。结果是从单个基因可以产生出选择性的RNA。

学习（Learning）

涉及概率的数据获取，无需记忆具体的例子。

学习算法（Learning algorithms）

让计算机能够学习的程序。

压缩（Compression）

将信息载体通过计算转换成更小的对象，并且压缩后的对象能够通过计算还原成原来较大的形式（解压）。

压缩深度（Compression depth）

数字对象的可逆压缩所需的计算资源（时间或空间）的数学度量。

氧化/还原（Oxidation/reduction）

电子在分子之间发生转移的化学反应。在氧化/还原反应中，如果一种反应物（参与反应的分子）被氧化，另一种则被还原。获得电子的反应物被还原，失去电子的反应物则被氧化。

遗传码（Genetic code）

DNA（和mRNA）中核苷酸的顺序与蛋白质中氨基酸的顺序的对应规则。mRNA的核苷酸每次读取3个，每3个核苷酸（1个密码子）对应于多肽中一个特定的氨基酸。参见密码子和mRNA。

遗传算法（Genetic algorithm）

模仿生物进化的“突变”和“重组”的计算机算法，属于进化算法的一种。

遗传性变异（Heritable variation）

在同种类的自然种群中总是存在基因差异。整个种群的这种差异称为遗传性变异。群体中每个成员的基因组都有微小的差异。参见基因组。

异或（XOR）

一种布尔函数，当且仅当输入有一个为1时结果为1，否则结果为0。参见布尔函数。

意向性立场（Intentional stance）

丹尼尔·丹尼特提出的术语，用来描述人们将人类动机的语言应用于自然界中的情形。例如：“鱼有尾巴是为了在水里游泳。”

意义（Meaning）

当一个对象或行为对某物或某人具有特定的影响，意义就出现了。意义意味着某个对象对于其他对象具有重要性。

有序（Order）

随机的对立面。有序表明模式或结构的存在。

与（AND）

一种布尔函数（参见布尔函数），当所有输入为1时输出值为1（真），否则为0（假）。

元胞自动机（Cellular automaton）

“元胞”组成的1维、2维或多维网格。元胞自动机根据规则周期性更新状态，新的状态取决于元胞本身以及相邻元胞的当前状态。

原核生物（Prokaryote）

基于原核细胞的生物。原核细胞比真核细胞简单，没有细胞核。细菌是最常见的原核生物。

原子（Atom）

与化学有关的最小物质单位。每种元素的原子都有特定数量的质子和电子。

折返（Reentry）

杰拉德·埃德尔曼用这个词来描述“脑区之间并行进行的信号动态交换，使得脑区在时间和空间上持续关联。”折返利用神经元群内部和之间的大规模并行往返连接组成的庞杂网络。埃德尔曼进一步将折返描述为一种大脑中持续的高层次选择形式。

真核生物（Eukaryote）

基于真核细胞的生物。真核细胞要比原核细胞复杂得多，包含大量内部结构，包括细胞核和线粒体。所有高等生物（植物、动物和真菌）都由真核细胞组成。

整体适应（Nonlocalized adaptation）

丹·阿什洛克提出的词，用来描述在基于计算机的协同进化系统中

观察到的普遍的能力提高。

指令（Instruction）

以某种方式组织，能使特定的事情发生或建立起某种结构的信息。

质粒（Plasmid）

从染色体DNA分离并且独立复制的DNA分子。质粒能携带各种蛋白质的基因编码。

质量（Mass）种子行（Seed row）

物质的基本属性。质量度量物体对加速的阻碍，也决定物体对其他质量物体的引力有多强。质量与重量密切相关，但不是专门针对地球。元胞自动机的第一行。参见元胞自动机。

重组（Recombination）

生物学和进化算法中使用的术语。在生物学中指的是两个DNA分子截断然后交互重连的过程。在计算机科学中有类似的意义，不过重组的是数据结构而不是DNA。重组也称为“交叉”。

主要组织相容性（Major histocompatibility, MHC）

一组基因及其编码的细胞表面蛋白质从原则上决定组织类型和移植相容性。MHC蛋白质在体内几乎所有细胞上都有，在T细胞对外来抗原的识别中扮演了重要角色。

转录（Transcription）

RNA合成。在分子遗传学中，转录指的是复制DNA序列形成RNA序列的过程。

转录因子（Transcription factor）

调控从DNA模板合成RNA的蛋白质。通常是通过与DNA结合并与RNA聚合酶相互作用实现。

转移RNA（transfer RNA, tRNA）

将氨基酸组合成核糖体并根据mRNA的核苷酸序列进行排列的一类RNA分子。

转译（Translation）

在分子遗传学中，转译指的是蛋白质的合成过程。

转座酶（Transposase）

负责从DNA大分子上剪切出一小段DNA并插入新位置的酶。

转座子（Transposon）

由转座酶从DNA大分子上剪切出来并插入新位置的一段DNA。转座子经常携带有编码转移它的转座酶所需的基因。如果是这种情况，转座子就有高度的独立性，在适当的条件下可以转移到新的细胞甚至不同的物种。

状态（State）

系统组分的特定的组合形态。

状态概率（State probability）

系统出现特定状态的可能性。

自催化集（Autocatalytic set）

斯图尔特·考夫曼用这个词描述化学反应组成的网络，其中一些化学成分是一些化学反应的催化剂。如果化学成分的种类足够多，并且其中有很大一部分是网络中化学反应的催化剂，则只要有适当的化学输入，网络就能一直自我维持下去。

自然选择（Natural selection）

在达尔文式进化中，自然决定哪些生物生存下来并繁育下一代。

自组织临界性（Selforganized criticality）

系统趋向介于稳定和不稳定之间的平衡状态。

自组装（Selfassembly）

多个部件无需外界帮助组合到一起形成结构的过程。这其中“没有组装者”，通常是热运动将部件移入位置。

最多1（One-max）

一个简单的算法问题，用程序将0/1串转化成同样长度的全1串，但没有明确告诉计算机如何识别1。如果允许做加法，进化算法可以轻松实现。

最佳击球点（Sweetspot）

击打棒球时，如果球棒的特定部位击中球，球会飞得更远，击球手的手也不会痛。球棒的这个部位称为最佳击球点。这个词用在其他背景下意指很大范围中一个有很好特性的小区域。

参考文献

Adams, Douglas. 1979.The Hitchhikers Guide to the Galaxy.London：Pan Books.

Adami, Christoph, Charles Ofria, and Travis C. Collier.2000.“Evolution of Biological Complexity.”Proceedings of the National Academy of Sciences 97：4463—4468.

Ashby, W. Ross.1956.An Introduction to Cybernetics.London：Chapman&Hall.

Ashlock, Daniel. 2006.Evolutionary Computation for Modeling and Optimization.New York：Springer

Ashlock, Daniel and John E. Mayfield.1998.“Acquisition of General Adaptive Features by Evolution.”In Evolutionary ProgrammingⅦ.V.W.Porto, N.Saravanan, D.Waagan, and A.E.Eiben eds.，75—84.London：Springer.

Ashlock, Daniel, Elizabeth Blankenship, and Jonathan Gandrud. 2003.“A Note on General Adaptation in Populations of Painting Robots.”In Proceedings of the 2003 Congress on Evolutionary Computation，46—53，IEEE Press.

Ashlock, Daniel and Brad Powers. 2004.“The Effect of Tag Recognition on Non-Local Adaptation.”In Proceedings of the 2004 Congress on Evolutionary Computation, vol 2.2045—2051.IEEE Press.

Ashlock, Daniel and Adam Sherk. 2005.“Non-Local Adaptation of Artificial Predators and Prey.”Proceedings of the 2005 Congress on Evolutionary Computation.41—48.IEEE Press.

Bak, Per. 1996.How Nature Works, The Science of Self-Organized

Criticality.New York：Copernicus.

Bear, Mark F.，Barry W. Conners, and Michael Paradiso.2007.Neuroscience，3rd ed.-

Baltimore：Lippincott, Williams and Wilkins.

Behe, Michael. 1996.Darwins'Black Box：The Biochemical Challenge to Evolution.

New York：Touchstone.

Beinhocker, Eric. 2006.The Origin of Wealth：Evolution, Complexity, and the Radical Remaking of Economics.Cambridge：Harvard Business School Press.

Bennett, Charles. 1988.“Logical Depth and Physical Complexity.”In The Universal Turing Machine：A Half Century Survey.Rolf Herken, ed.，207—236.Wien：Springer Verlag.

Borges, Jorge Luis 1962.“The Library of Babel，”51—58，In Labyrinths. New York：New ；Directions.

Bray, Dennis. 2009.Wetware, a Computer in Every Living Cell.New Haven：Yale University Press.

Campbell, Donald T. 1974.“Evolutionary Epistemology.”In The Philosophy of Karl Popper, P.A.Schlipp, ed.，Chapter 12，412—463.Chicago：Open Court.

Campbell, Neil A. and Jane B.Reese.2008.Biology，8th ed.San Francisco：Benjamin-Cummings.

Carroll, Sean. 2005.Endless Forms Most Beautiful, the New Science of Evo Devo and the Making of the Animal Kingdom.New York：W.W.Norton.

Compton, Arthur H. 1940.The Human Meaning of Science.Chapel

Hill：University of North Carolina Press.

Compton, Arthur H. 1935.The Freedom of Man.New Haven：Yale University Press.

Cover, Thomas M. and Joy A.Thomas.1991.Elements of Information Theory.New York：Wiley.

Cyran, S. A.，A.M.Buchsbaum, K.L.Reddy, M.C.Lin, N.R.Glossop, P.E.Hardin, M.W.Young, R.V.Storti, and J.Blau.2003.“vrille, Pdp1，and dClock Form a Second Feedback Loop in the Drosophila Circadian Clock.”Cell 112：329—341.

Cziko, Gary. 1995.Without Miracles.Cambridge：MIT Press.

Darwin, Charles. 1839.The Voyage of the Beagle.London：Colburn.

Darwin, Charles. 1859.On the Origin of Species by Means of Natural Selection.London：John Murray.

Darwin, Charles. 1871.The Descent of Man, and Selection in Relation to Sex.London：John Murray.

Davies, Paul and Niels Henrik Gregersen, eds. 2010.Information and the Nature of Reality.Cambridge：Cambridge University Press.

Dawkins, Richard. 1976.The Selfish Gene.Oxford：Oxford University Press.

Dawkins, Richard. 1986.The Blind Watchmaker.New York：W.W.Norton，

Dawkins, Richard. 1996.Climbing Mount Improbable.New York：W.W.Norton.

Dawkins, Richard. 2004.The Ancestors Tale：A Pilgrimage to the Dawn of Evolution.Boston：Houghton Mifflin.

Dennett, Daniel. 1995.Darwins'Dangerous Idea：Evolution and the Meanings of Life.New York：Simon&Shuster.

Dennett, Daniel. 2003.Freedom Evolves.New York：Viking.

Devaney, Robert L. 1990.Chaos, Fractals, and Dynamics, Computer Experiments in Mathematics.Boston：Addison-Wesley.

Dorit, Robert. 1997.“A Review of Darwins'Black Box：The Biochemical Challenge to Evolution.”In American Scientist 85；474—475.

Doty, David. 2004.“Non-Local Evolutionary Adaptation in Gridplants.”Proceedings of the 2004 Congress on Evolutionary Computation vol 2.1602—1609.IEEE Press.

Dyson, George B. 1997.Darwin Among the Machines：The Evolution of Global Intelligence：New York：Perseus Books.

Edelman, Gerald. 1987.Neural Darwinism：A Theory of Neuronal Group Selection.New York：Basic Books.

Edelman, Gerald and Giulio Tononi. 2000.A Universe of Consciousness：How Matter Becomes Imagination.New York：Basic Books.

Eigler, D. M.and E.K Schweizer，.1990.“Positioning Single Atoms with a Scanning Tunneling Microscope.”Nature 344：524—526.

Erwin, Douglas H.，Marc Laflamme, Sarah M. Tweedt, Erik A.Sperling, Davide.Pisani.2011.“The Cambrian Conundrum：Early Divergence and Later Ecology Success in the Early History of Animals.”Science 334：1091—1097.

Fennimore, A. M.，T.D.Yuzvinsky, Wei-Qiang Han, M.S.Fuhrer, J.Cumings, and A.Zettl.2003.“Rotational Actuators Based on Carbon Nanotubes.”Nature 424：408—410.

Friston, Karl. 2005.“A Theory of Cortical Responses.”Philosophical Transactions of the Royal Society B：Biological Sciences 360：815—836.

Friston, Karl. 2010.“The Free Energy Principle.”Nature Reviews Neuroscience2：127—138.

Gilbert, Scott F. 2006.Developmental Biology，8th ed.Sunderland：Sinauer.

Goldberg, David. 1989.Genetic Algorithms in Search Optimization and Machine Learning.Boston：Addison Wesley.

Goldstein, Bernard, Michele Demak, Mary Northridge, and Daniel Wartenberg. 1992.“Risk to Groundlings of Death Due to Airplane Accidents：A Risk Communication Tool.”Risk Analysis，12：339—341.

Gordon, Peter. 2004.“Numerical Cognition Without Words：Evidence from Amazonia.”Science 306：496—499.

Gould, Steven J. 1989.Wonderful Life：The Burgess Shale and the Nature of History.New York：W.W.Norton.

Gould, Steven J. 1996.Full House：The Spread of Excellence from Plato to Darwin.New York：Three Rivers.

Greene, Brian. 2003.The Elegant Universe：Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory.New York：W.W.Norton.

Gregory, T. R.2004.“Macroevolution, Hierarchy Theory, and the C-Value Enigma.”Paleobiology 30：179—202.

Heylighen, Francis. 1999.“The Growth of Structural and Functional Complexity During Evolution.”In The Evolution of Complexity：The Violet Book of‘Einstein Meets Magritte'（Einstein Meets Magritte：An Interdisciplinary Reflection on Science, Nature, Art, Human Action and Society）.Francis Heylighen, Johan Bollan, and Alexander Riegler eds.，17

—44.Dordrecht：Kluwer.

Holland, John. 1992.Adaptation in Natural and Artificial Systems：An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence，2nd ed.Cambridge：MIT Press.

Holland, John. 1995.Hidden Order：How Adaption Builds Complexity.New York：Perseus Books.

Holland, John. 1998.Emergence：From Chaos to Order.New York：Perseus Books.

Hull, David. 1988.Science as a Process：An Evolutionary Account of the Social and Conceptual Development of Science.Chicago：University of Chicago Press.

Inlay, Matt. 2002.“Evolving Immunity, A Response to Chapter 6 of Darwin's Black Box.”网址http：//www.talkdesign.org/faqs/Evolving_Immunity.html.2013年1月15日访问.

Janeway, Charles, Mark Shlomchk, Paul Travers, and Mark Walport. 2004.Immunobiology.New York：Garland Science.

Juedes, David W.，James I. Lathrop, and Jack H.Lutz.1994.“Computational Depth and Reducibility.”Theoretical Computer Science 132（2）：37—70.

Kandel, Eric. 2007.In Search of Memory：The Emergence of a New Science on Mind.New York：W.W.Norton.

Kauffman, Stuart. 1993.Origins of Order：Self-Organization and Selection in Evolution.Oxford：Oxford University Press.

Kauffman, Stuart. 1995.At Home in the Universe：The Search for Laws of Self-Organization and Complexity.Oxford：Oxford University Press.

Kauffman, Stuart. 2008.Reinventing the Sacred：A New View of Science, Reason, and Religion.New York：Basic Books.

Knill, David C. and Alexandre Pouget.2004.“The Bayesian Brain：The Role of Uncertainty in Neural Coding and Computation.”TRENDS in Neuroscience 27：712—719.

Knoll, Andrew H. 2003.Life on a Young Planet, the First Three Billion Years of Evolution on Earth.Princeton：Princeton University Press.

Koschmieder, E. L.1993.Bénard Cells and Taylor Vortices.Cambridge：Cambridge University Press.

Kuhn, Thomas. 1962.The Structure of Scientific Revolutions.Chicago：University of Chicago Press.

Kurzweil, Ray. 2005.The Singularity Is Near.London：Penguin Books.

Lathrop, James. 1997.PhD Dissertation.Computing and Evolving Variants of Computational Depth.Ames：Iowa State University.

Lathrop, James. 1997“.Compression Depth and Genetic Programs.”In Proceedings of the Second Annual Conference on Genetic Programming，370—378.San Francisco：Morgan Kaufmann.

Lee, Tai Sing and David Mumford. 2003.“Hierarchial Bayesian Inference in the Visual Cortex.”Journal of the Optical Society of America A 20：1434—1448.

Leiman, P. G.，S.Kanamaru, V.V.Mesyanzhinof, F.Arisaka, and M.G.Rossmann.2003.“Structure and Morphogenesis of Bacteriophage T4.”Cellular and Molecular Life Sciences 60：2356—2370.

Lincoln, Tracy A. and Gerald F.Joyce.2009.“Self-Sustained Replication of an RNA Enzyme.”Science 323：1229—1232.

Lloyd, Seth. 2006.Programming the Universe：A Quantum Computer

Scientist Takes on the Cosmos.New York：Alfred A.Knopf.

McShea, Daniel W. and Robert N.Brandon.2010.Biologys'First Law：The Tendency for Diversity&Complexity to Increase in Evolutionary Systems.Chicago：University of Chicago Press.

McShea, Daniel W. 1991.“Complexity and Evolution；What Everybody Knows.”Biology and Philosophy 6：303—324.

Morowitz, Harold J.，Vijaysarathy Srinivasan, Shelley Copley, and Eric Smith. 2005.“The Simplest Enzyme Revisited, the Chicken and Egg Argument Solved.”Complexity10：12—13.

Mossel, Elchanan and Mike Steel. 2004.“Random Biochemical Networks：The Probability of Self Sustaining Autocatalysis.”Journal of Theoretical Biology 233：327—336.

Paley, William. 1802.Natural Theology—or Evidences of the Existence and Attributes ofthe Deity Collected from the Appearances of Nature.London：Richard Faulder.

Pinker, Steven. 1997.How the Mind Works.New York：W.W.Norton.

Popper, Karl“Of Clocks and Clouds.”1972. Chapter 6 inObjective Knowledge：An Evolutionary Approach.Oxford：Oxford University Press.

Popper, Karl. 1968.The Logic of Scientific Discovery.New York：Harper.

Powner M. W.，B.Gerland, and J.D.Sutherland.2009.“Synthesis of activated Pyrimidine Ribonucleotidesin Prebiotically Plausible Conditions.”Nature 459：239—242.

Purvis, Dale, George J. Augustine, David Fitzpatrick, William C.Hall, Anthony-Samuel La Mantia, and Leonard E White, eds.2012.Neuroscience.5th ed.Sunderland：Sinauer Associates.

Rode, Bernd M.，Daniel Fitz, and Thomas Jakschitz. 2007.“The First Steps of Chemical Evolution Towards the Origin of Life.”Chemistry&Biodiversity 4：2674—2701.Zurich：Verlag Helvetica Chemica Acta.

Rothemund, Paul W. K.，Nick Papadakis, and Eric Winfree.2004.“Algorithmic Self-Assembly of DNA Sierpinski Triangles.”PLOS Biology 2：2041—2052.

Ruse, Michael. 2003.Darwin and Design：Does Evolution Have a Purpose?Cambridge：Harvard University Press.

Schopf, William, ed. 2002.Lifes'Origin.Berkeley：University of California Press.

Scientific American. 2002.Understanding Artificial Intelligence.New York：Scientific American.

Sepkoski, J. J.1984.“A Kinetic Model of Phanerozoic Taxonomic Diversity.Ⅲ.Post-Paleozoic Families and Mass Extinctions.”Paleobiology 10：246—267.

Shapiro, Robert. 1987.Origins, A Skeptics'Guide to the Creation of Life on Earth.New York：Summit Books.

Signor, P. W.1994.“Biodiversity in Geological Time.”American Zoologist 34：23—32.

Smith, Adam. 1776.An Inquiry into the Nature and Causes of the Wealth of Nations.London：William Strahan.

Smolin, Lee. 1997.The Life of the Cosmos.Oxford：Oxford University Press.

Valentine, James W.，Allen G. Collins, and C.Porter Meyer.1994.“Morphological Complexity Increase in Metazoans.”Paleobiology 20：131—142.

Van Valen, Leigh. 1973.“A New Evolutionary Law.”Evolutionary Theory 1：1—30.

Wilson, David Sloan. 2002.Darwin's Cathedral, Evolution, Religion, and the Nature of Society.Chicago：Chicago University Press.

Wright, Sewall. 1932.“The Rolesof Mutation, Inbreeding, Crossbreeding, and Selection in Evolution.”In Proceedings of the Sixth International Congress on Genetics，356—366.

Wolfram, Stephen. 2002.A New Kind of Science.Champaign：Wolfram Media.

Wolpert, Lewis, Jim Smith, Tom Jessell, Peter Lawrence, Elizabeth Robertson. 2006.Principles of Development，3rd ed.Oxford：Oxford University Press.

译名表

专用名词

《盲眼钟表匠》|*The Blind Watchmaker*

《爱丽丝镜中奇遇记》|*ThroughtheLooking-Glass*

《达尔文的大教堂》|*Darwin’s Cathedral*

《科学革命的结构》|*TheStructureofScientifc Revolutions*

《人类的由来》|*Descent of Man*

《小猎犬的远航》|*Voyage of the Beagle*

《银河系搭车客指南》|*TheHitchhiker’sGuideto the Galaxy*

《自私的基因》|*The Selfsh Gene*

阿肯色大学|University of Arkansas

爱荷华州立大学|Iowa State University

哥伦比亚大学|Columbia University

圭尔夫大学|University of Guelph

华盛顿大学|Washington University

加州大学伯克利分校|University of California at Berkeley

加州理工学院|Caltech

卡内基·梅隆大学|Carnegie Mellon University

伦敦大学学院|University College London

麻省理工学院|MIT

曼切斯特大学|University of Manchester

美国国家科学基金|National Science Foundation

美国国家卫生研究院|National Institutes of Health

密歇根大学|University of Michigan

牛津大学|Oxford University

普林斯顿大学|Princeton University

沃尔夫勒姆研究公司|Wolfram Research, Inc

芝加哥大学|University of Chicago

人名

理查德·道金斯|Richard Dawkins

阿兰·阿瑟利|Alan Atherly

阿瑟·康普顿|Arthur H.Compton

阿斯拓洛·特勒|Astro Teller

埃里克·拜因霍克尔|Eric Beinhocker

埃里克·温弗里|Erik Winfree

埃米·迪坎普|Amy D’Camp

奥德丽·吉福德|Audrey Gifford

彼得·丹宁|Peter Denning

彼得·普雷斯科特|Peter Prescott

玻尔兹曼|Ludwig Boltzmann

布兰登·霍尔特|Brandon Holt

布里奇特·弗兰纳里-麦科伊|Bridget Flannery-McCoy

大卫·赫尔|David Hull

大卫·威尔逊|David Sloan Wilson

丹·阿什洛克|Dan Ashlock

丹尼尔·丹尼特|Daniel Dennett

丹尼尔·麦克谢伊|Daniel McShea

道格·芬尼莫尔|Doug Finnemore

道格拉斯·亚当斯|Douglas Adams

杜安·恩格尔|Duane Enger

戈登·鲍恩|Gordon Bowen

戈登·摩尔|Gordon Moore

吉布斯|Willard Gibbs

吉姆·琼斯|Jim Jones

加里·霍华德|GaryHoward

杰克·卢茨|Jack Lutz

杰拉尔德·埃德尔曼|Gerald Edelman

卡尔·波普尔|Karl Popper

卡尔·弗里斯通|Karl Friston

康德|Auguste Comte

柯尔莫哥洛夫|Andrei Kolmogorov

科施米德尔|E.L.Koschmeider

克雷格·马爹利|Craig Martell

克雷格·文特尔|Craig Venter

雷伊·克兹维尔|Ray Kurzweil

利·范瓦伦|Leigh Van Valen

利·特斯特法新|Leigh Testfatsion

罗斯·阿什比|W.Ross Ashby

马丁·斯波尔丁|Martin Spalding

马特·帕蒂兹|Matt Patitz

梅格·皮珀布林克|Meg Pieperbrink

米根·赫瑟尔顿|Meaghan Hetherton

牛顿|Isaac Newton

帕特里克·菲茨杰拉德|Patrick Fitzgerald

乔纳森·甘拉德|Jonathan Gandrud

乔治·布尔|George Boole

塞思·劳埃德|Seth Lloyd

史蒂芬·卡瓦莱尔|Steve Kawaler

斯宾塞|Herbert Spencer

斯蒂芬·古尔德|Stephen J.Gould

斯蒂芬·图尔明|Stephen Toumlin

斯蒂芬·沃尔弗拉姆|Stephan Wolfram

斯科特·麦克勒|Scott Mackler

斯科特·佩利|Scott Pelley

斯图尔特·考夫曼|Stuart Kauffman

塔莎·奥布林|Tasha Obrin

唐·坂口|Don Sakaguchi

唐纳德·坎贝尔|Donald Campbell

托马斯·贝叶斯|Thomas Bayes

托马斯·贝叶斯|Thomas Bayes

托马斯·库恩|Thomas Kuhn

威廉·佩利|William Paley

香农|Claude Shannon

肖恩·西贝尔斯基|Sean Siberski

伊丽莎白·布兰肯希普|Elizabeth Blankenship

伊桑·斯科特|Ethan Scott

约翰·惠勒|John Wheeler

约翰·霍兰|John Holland

约瑟夫·史密斯|Joseph Smith

詹姆士·沃森|James Watson

詹姆斯·莱思罗普|James Lathrop

詹姆斯·莫罗|James Morrow

第一推动丛书: 综合系列

皇帝新脑

[英] 罗杰·彭罗斯 著

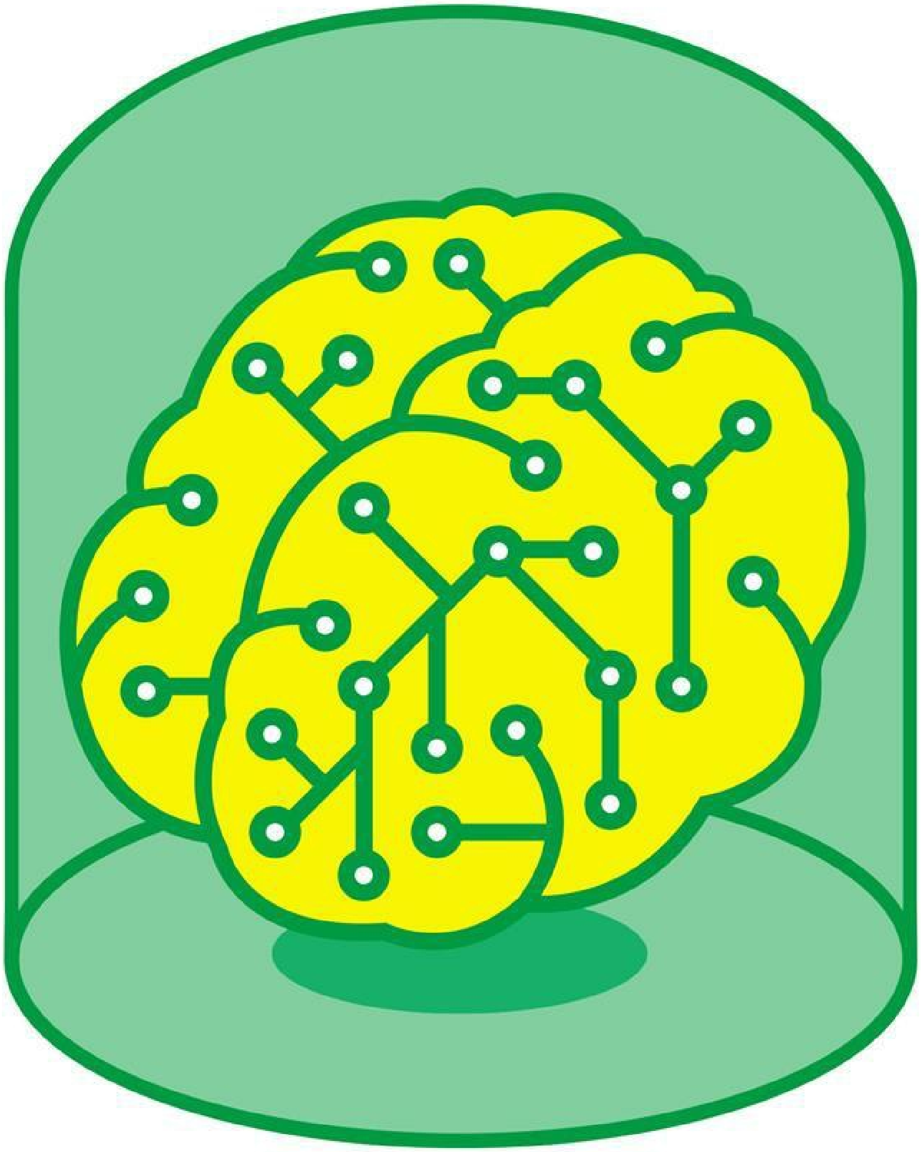
许明贤 吴忠超 译

The Polytechnique Series

The Emperor's New Mind

Roger Penrose

THE FIRST MOVER



湖南科学技术出版社

版权信息

皇帝新脑

著者：[英] 罗杰·彭罗斯

译者：许明贤　吴忠超

责任编辑：李永平　吴炜　戴涛　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794444

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
题献
敬启读者
译者序
前言
感谢
序言
第1章 电脑能有精神吗
第2章 算法和图灵机
第3章 数学和实在
第4章 真理、证明和洞察
第5章 经典世界
第6章 量子魔术和量子神秘
第7章 宇宙论和时间箭头
第8章 量子引力的寻求
第9章 真实头脑和模型头脑
第10章 精神物理的寻求
注释
跋

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

题献

谨将此书献给未能活到见证它问世的我亲爱的母亲。

敬启读者

我在本书的许多地方引用数学公式，而毫不在乎时常听到的警告——放进去的每条公式都会把我的读者数目减半。如果你是这样的一位对数学公式恐惧的读者（大多数人都是这样的），那么我就介绍一种当这种可恨的一行出现时自己通常采用的步骤。大体就是完全不理睬这一行，而跳到正文的下一行去！但也并非完全如此；人们要仔细地推敲这一可怜的公式，而不仅是做表面上的理解，然后再继续前进。过了一阵，如果你又重新充满自信，则可回到刚才忽略了的公式，努力抓住一些显著的特征。正文本身也许可帮助你了解什么是关键的，什么东西被忽略后并没有什么影响。如果做不到这些，也不必担心，干脆不理该公式就是了。

译者序

许明贤　吴忠超

1992年1月9日

纽约　长岛

牛津大学的罗杰·彭罗斯的《皇帝新脑》一书的出版是国际书界的一件大事。剑桥大学前年曾为它专门举行了一次学术会议。

这本洋洋大观的贯穿了计算机科学、数学、物理学、宇宙学、神经和精神科学以及哲学的巨著，体现了作者向哲学上最大的问题之一“精神——身体关系”挑战的大无畏精神。

迄今为止的科学基本上都可纳入形而下学的范畴，而这本书可认为是首次对形而上学进行的严肃尝试。历史上曾重复地出现过还原主义的思潮，最近代的便是人工智能专家的断言：电脑最终能代替人脑甚至超过人脑。彭罗斯的论断却是：正如皇帝没有穿衣服一样，电脑并没有头脑。电脑具有智慧吗？人们的共识是用通过图灵检验来定义智慧。彭罗斯认为要制造出满意地通过这种检验的机器还是非常遥远的事。即使它真的通过了，我们还是不能断定其真有理解力，西尔勒中文屋子的理想实验强有力地表明，用图灵检验来定义智慧还是远远不够充分的。

希尔伯特曾经有过一个非常宏伟的计划，一旦公理和步骤法则给定，一切真理都应该能被推导出来。著名的哥德尔定理使这个计划的宏图化为泡影。以算法来获取真理的手段是非常受局限的，在任何一个形式系统中总存在不能由公理和步骤法则证明或证伪的正确的命题。康托尔关于无理数集合的不可数性、罗素集论的理发师悖论、哥德尔定理以及电脑停机问题都是一脉相承地沿用了康托尔对角线法而给予证明的。一言以蔽之，世界万花筒般的复杂性不可能用可数的算法步骤来穷尽。

灵感和直觉在发现真理方面比逻辑推导重要得多。彭罗斯和柏拉图相认同，发现真理是精神和数学观念的柏拉图世界进行接触。正如询问宇宙在大爆炸之前是什么样子的问题是没有意义的一样，柏拉图世界是超越时空的，具有“遗世独立”的品格。柏拉图世界至少和物理世界一样地具有实在性甚至两者是合二为一的。他认为著名的芒德布罗集一定是栖息在这个世界中，否则的话何以这么美丽呢？真可谓：“此曲只应天上有，人间能得几回闻？”

牛顿力学、麦克斯韦电磁学、爱因斯坦相对论和量子力学给人类带来了神速的技术进步，在使人们充满了自信心的同时也给套上了宿命论的枷锁。我们宇宙的一切都已完全为第一推动所决定。过去人们将第一推动归于上帝，而量子宇宙论却把第一推动也都摒除了，宇宙在时空上是有限无界的！这肯定是自以为具有自由意志的人类所不能忍受的，什么人愿意生活在这种宇宙中呢？远离平衡态的热力学耗散结构也许是生命现象的雏形，动力系统的不稳定性导致混沌之中又隐含着新的秩序，这些是对理解生命的努力，也是半个世纪前人们始料不及的，但这还不是形而上学的精神。彭罗斯猜测，宇宙也许的确是宿命论的，但同时是不可计算的。我们的宇宙究竟有自由意志的存身之所吗？

人类智慧的最伟大工程之一是爱因斯坦的统一场论。其主要困难在于量子力学和相对论之间的不协调。绝大多数物理学家都责备相对论，认为广义相对论只是一种唯象的理论。其实，就理论的美丽和经济性而言，相对论是远远地比量子力学优越。前者是人类智慧的产物，而后者是人们不得不接受的规则。量子力学的解释中仍有许多问题，譬如波函数的坍缩、薛定谔猫佯谬和爱因斯坦——玻多尔斯基——罗森“矛盾”。这些困难也许在超越过它们的量子引力中可以得到解决。广义相对论的美丽和经济性体现在非线性之中，彭罗斯曾提出过非线性引力子的概念，这是从他早年对引力波碰撞的研究中得到启发的。他猜测到，发生量子波函数的坍缩的判据在于其引力效应超过单引力子的阈值。

彭罗斯镶嵌是除了芒德布罗集之外的对柏拉图观念存在性的有力支持。这两个例子的共同性是它们的发现和近代科学的进展基本无关。准晶体的五重对称性是这种镶嵌的三维体现。彭罗斯猜测，准晶体生长的神经元行为既涉及单引力子判据又涉及量子引力的非定域性。

时间及其方向也许是意识的最大秘密。彭罗斯提出了外尔曲率猜

测，宇宙的引力熵由外尔曲率来度量，而在大爆炸奇点处它必须为零，可惜迄今连这种关系的表达式都还没有找到，也许它必须是非定域的。他认为时间流逝的方向是由此衍生而来的。

原子时间、生物时间和宇宙时间以及时间箭头只不过是对时间概念的粗糙近似。爱因斯坦——玻多尔斯基——罗森“矛盾”表明波函数坍缩是和狭义相对论的定域性以及因果性相矛盾，更遑论广义相对论了。在精神现象中，甚至时序都发生混乱，在灵感、直觉过程中或者在与柏拉图世界接触时似乎时间被不可思议地压缩了，它们甚至不发生在时间里。我们在洋洋自得的同时，又发现科学理论的成就还是这么贫弱。要完全弄清时间的含义得有待于量子引力的成功，这也是推动精神物理发展的关键。

彭罗斯对引力物理有过许多重要贡献，他（和霍金一道）证明了广义相对论的奇点的不可避免性，提出了黑洞的捕获面，以及克尔黑洞的能层概念。他发明了研究时空拓扑结构的主要工具即彭罗斯图。他对类空、类时和零无穷的阐释使引力辐射的图像更具形象。他把旋量引进引力物理，使辐射问题的研究更新，这就是纽曼——彭罗斯形式，在此框架中他证明了剥皮定理，即向无穷远辐射的引力可按照其衰弱方式被分成4个层次（电磁波只有两个层次）。

本书充满了许多猜测，正如历史上的许多猜测的命运一样，一些会存活，另一些会被淘汰。不管它们的命运如何，这正是当代思想家、哲学家和科学家必须去做正面冲突的问题。本书的字里行间充满了作者探索真理的灵感和激情。译者历经一个寒暑的辛苦，终于把这个译本奉献在读者面前。但愿在浏览此书之时，会有王献之行走于山阴道上目不暇接之感。人们在忙碌于都市生活之余，抽空到兰亭一游不也是件赏心乐事吗？

前言

马丁·伽特纳

许多伟大的数学家和物理学家觉得，要写一本外行能理解的书，如果不是不可能的话，也是非常困难的。直到今年，人们也许还认为，罗杰·彭罗斯，这位世界上最博学和最有创见的数学物理学家之一，也属于这个范畴之内。我辈读过他的非专业性的文章和讲演，稍微了解一些底细。尽管如此，当发现彭罗斯在他的研究之余花费大量时间为见多识广的外行写下了这样美妙无比的书时，人们的确感到惊喜。我相信，该书会成为一部经典。

虽然彭罗斯的著述广泛地涉及相对论、量子力学和宇宙论，其关心的焦点乃是哲学家所谓的“精神——身体问题”。几十年来，人工智能专家尽力说服我们，再有一两个世纪的时间（有些人已把这些时间缩短到50年！）电脑就能做到人脑所能做的一切。他们因为受年轻时读到的科学幻想的刺激，而坚信我们的精神只不过是“肉体的电脑”（正如马文·闵斯基曾经提出过的）。他们想当然地认为，当电子机器人的算法行为变得足够复杂时，痛苦和快乐、对美丽和幽默的鉴赏、意识和自由意志就会自然地涌现出来。

有些科学哲学家（最著名者为约翰·西尔勒，他的大名鼎鼎的中文屋子的理想实验为彭罗斯所深入讨论）强烈地反对这种看法。对他们来说，电脑和用轮子、杠杆或任何传递信号的东西运行的机械计算机并没有什么本质的不同（人们可用滚动的弹子或通过管道流动的水流制造计算机）。因为电流通过导线比其他能量形式（除了光）走得更快，它就能比机械计算机更快地摆弄信号，并因此能承担庞大复杂的任务。但是，一台电脑是否以一种比算盘更优越的方式“理解”它的所作所为呢？是的，现代电脑能以大师的风度下棋。它们是否比一群电脑迷曾经用积木搭成的方格游戏机（一种西文的初级游戏）对游戏“理解”得更好些？

彭罗斯的书是迄今为止对强人工智能的最猛烈的攻击。几个世纪以

来，人们就一直反对还原主义者关于精神只不过是已知物理定律操纵的机器的宣称。但是，因为彭罗斯凭借从前的作者不能获知的资讯，所以他的攻击更加令人信服。从这本书可以看出，彭罗斯不仅是一位数学物理学家，而且是一位第一流的哲学家，他毫无畏惧地和当代哲学家斥之为无稽的问题进行搏斗。

彭罗斯还不顾一小群物理学家的越来越强烈的否定，敢于认可坚定的实体主义。不仅宇宙是“外在的”，而且数学真理自身也有其神秘的独立性和永恒性。正如牛顿和爱因斯坦那样，彭罗斯对物理世界和纯粹数学的柏拉图实体极其谦恭和敬畏。杰出的数论学家保罗·厄多斯的口头禅是，所有最好的证明都记载在“上帝的书”上，数学家偶尔被允许去瞥见一页半纸。彭罗斯相信，当一位物理学家或者数学家经历一次突然的“惊喜”的洞察，这不仅是“由复杂计算作出”的某种东西，而是精神在一瞬间和客观真理进行了接触。他感到惊讶，莫非柏拉图世界和物理世界（物理学家已将其融入数学之中）真的是合二为一？

彭罗斯用了不少篇幅论及以其发现者贝内特·芒德布罗命名的芒德布罗集的著名的类分数维结构。虽然其局部放大在统计的意义上是自相似的，它的无限地盘旋的模式却以不可预见的方式不断地改变。彭罗斯（和我一样）觉得，若有人不认为这一奇异的结构不像喜马拉雅山那样是“外在的”，而且有待人们像探险丛林那样去勘探，那真是不可理喻。

彭罗斯是数量不断增加的一伙物理学家的一员，认为当爱因斯坦说他的“小指”告诉他量子力学是不完备时，他并非顽冥不化或昏头昏脑。彭罗斯为了支持这一争论，把你指引向涵盖众多课题的旅途，诸如复数、图灵机、复杂性理论、哥德尔的不完备性、相空间、希尔伯特空间、黑洞、白洞、霍金辐射、熵、脑结构以及许多当代研究的核心问题。狗和猫对其自身有“意识”吗？传递物质的机器可能在理论上把一个人如同在电视系列片《星际旅行》中那样，把宇航员从上往下地扫描的办法从一处向另一处运送吗？进化在意识的产生中发现了什么存活的价值？是否存在超越量子力学的一种水平，它为时间的方向以及左右之间的差别刻上烙印？量子力学的定律，也许甚至更高深的定律，是否对精神现象具有根本的作用？

彭罗斯对上述的最后两个问题的回答是“是”。他的著名的“扭量”理论——在作为空间——时间基础的高维复空间中运算的抽象的几何对象

——因为过于专业化而不能被包括在此书之中。它是彭罗斯20多年对比量子力学的场和粒子更深刻的领域进行探索的结果。在他对理论的4种分类，即超等、有用、尝试和误导之中，彭罗斯谦虚地把扭量理论和现在激烈争论的超弦以及其他大统一方案一道归于尝试类中。

彭罗斯从1973年起担任牛津大学的罗斯·玻勒数学教授。这个头衔对他甚为适合。因为W.W.罗斯·玻勒不仅是一位著名的数学家，还是一位业余魔术家。对数学游戏的强烈兴趣使他写下该领域的英文经典著作《数学游戏及漫笔》。彭罗斯和玻勒一样地热心于游戏。彭罗斯在年轻时发现了一种称为“三杆”的“不可能物体”（一个不可能物体是由于其自相矛盾而不能存在的立体形态的图画）。他和他的父亲列奥尼，一位遗传学家，把三杆转变成彭罗斯楼梯，毛里兹·伊歇把它用于两幅众所周知的石板画《升降》和《瀑布》之中。有一天彭罗斯躺在床上，他在“一阵狂热”之后摹想到四维空间中的不可能物体。他说，它是这样一种东西，甚至一个四维空间的生物遇到它的话也会惊叫：“天哪，这是什么东西？”

20世纪60年代，当他和朋友史蒂芬·霍金合作研究宇宙论时，作出了也许是他最著名的发现。如果相对性理论“一直下去”都是成立的，那么在物理学定律不再适用的每一黑洞里必须有一奇点。不过，甚至使这一成就黯然失色的是他近年的另一项成就：彭罗斯只用两种形状的花砖就能以伊歇镶嵌的方法把平面铺满，但是这种镶嵌只能采取非周期性的形式。（你们可在拙著《彭罗斯镶嵌》中见识到有关这些讨人喜欢的形状。）与其说他发明了它们，不如说发现了它们，当时一点也没预料到它们有何用场。当人们发觉，他的镶嵌的三维形式是物体的奇异的新形态基元时，不禁大为惊奇。现代晶体学最活跃的研究领域便是探讨这类“准晶体”。这也是好玩的数学找到预想不到应用的现代富有戏剧性的事件之一。

彭罗斯在数学和物理上的成就——我只能触及一小部分——源于他毕生对“存在”的神秘和美丽保持好奇之心。他的小指头告诉他，人脑不仅仅是小导线和开关的集合。他的序言和跋中的“亚当”一部分是知觉生命的缓慢进化的意识曙光的象征。依我看来，他也就是彭罗斯——坐在离开人工智能领导者第三排的地方——敢于直言人工智能的皇帝没有穿衣服。彭罗斯的许多看法都富有幽默感，但这件事情绝不是闹着玩的。

感谢

我在著作此书时，曾得到过许多人的各种帮助，在此谨表感谢。尤其是那些（特别是参与我所观看过的英国广播公司电视节目的）强人工智能的提倡者所表达的如此极端的人工智能的观点，在多年以前刺激了我着手这一规则。（然而，如果我早知道要完成此书竟要如此辛苦，恐怕当初就不敢开始！）许多人细读了手稿的小部分并提供了不少改进的建议，我对他们也表示谢意，他们是：托比·贝利、大卫·德义奇（他还检验了我的图灵机编号，对我特别有用）、斯图亚特·罕蒙帕希尔、詹姆·哈特尔、拉恩·休许斯顿、安古斯·马金泰、玛丽·詹·莫瓦特、崔斯坦·尼丹姆、特得·纽曼、埃利克·彭罗斯、托比·彭罗斯、沃尔夫冈·林得勒、恩格尔伯特·叙金和邓尼斯·西阿玛。我尤其欣赏克利斯托弗·彭罗斯为我提供的芒德布罗集的细节以及约纳逊·彭罗斯提供的弈棋电脑的有用信息。我特别感谢柯灵·伯勒克摩尔、埃利希·哈斯和大卫·胡贝尔，他们为我审阅了第9章，对于该章的领域我只能算是一个门外汉。正如我所致谢的其他人那样，本书的任何错误与他们无关。我感谢国家基金会的支持，其合同号码为DMS84—05644、DMS86—06488（莱斯大学，休斯敦，本书的一部分就是根据在该校作的演讲而写成的）以及PHY86—12424（希拉库斯大学，在该校进行了关于量子力学的有价值的讨论）。我还十分感谢马丁·伽特纳，他极其慷慨地为本书写前言以及提供一些具体的评论。我最感谢我亲爱的瓦尼莎，她对好几章进行了细致的批评，在文献上提供了许多帮助，在我最使人不能忍受时容忍我，她给了我极其需要的深挚的爱情和支持。

FIGURE ACKNOWLEDGEMENTS

THE PUBLISHERS EITHER have sought or are grateful to the following for permission to reproduce illustration material.

Figs 4. 6 and 4.9 from D.A.Klarner（ed.），*Themathematical Gardner*（Wadsworth International，1981）.

Fig. 4.7 from B.Grünbaum and G.C.Shephard，*Tilingsandpatterns*

（W.H.Freeman，1987）.Copyright©1987 by W.H.Freeman and Company.Used by permission.

Fig. 4.10 from K.Chandrasekharan，*HermannWeyl* 1885—1985（Springer，1986）.

Figs. 4.11 and 10.3 from pentaplexity：a class of non-periodictilings of the plane.*The Mathematical Intelligencer*，2，32—37（Springer，1979）.

Fig. 4.12 from H.S.M.Coxeter, M.Emmer, R.Penrose, and M.L.Teuber（eds），*M.C.Escher*：*Art and science*（North-Holland，1986）.

Fig. 5.2©1989 M.C.Escher Heirs/Cordon Ar-Baarn-Holland.

Fig. 10.4 from*JournalofMaterialsResearch*，2，1—4（Materials Research Society，1987）.

All other figures（including 4. 10 and 4.12）by the author.

序言

大会堂里有一个盛大的集会，标志着新的“超子”电脑的诞生。总统波罗刚刚结束了他的开幕词。他很高兴：他并不很喜欢这样的场合，对电脑也是一窍不通，只知道这种电脑即将为他赢得很多时间。制造商们向他保证，在这种电脑的诸多功能中，它还能代替他为那些他觉得如此厌倦的棘手的国家问题做决策。想到花费在它上面的金钱的数量，这种事最好是真的。他期待着能够在他那豪华的私人高尔夫球场上享受玩上许多小时高尔夫球的快乐——这是在他这小国家里所剩下为数不多的一块有相当面积的绿地。

亚当觉得置身于那些出席这一开幕典礼的人们之中不胜荣幸。他坐在第三排，两排前面坐着他的母亲：一个参加设计超子电脑的主要技术人员。凑巧的是，他的父亲也在那个场合——不过并没有得到邀请，现正在大厅后面被安保人员团团围住。在最后一分钟，亚当的父亲仍试图炸毁这台电脑。作为一小群灵魂意识委员会边缘活动分子的自命的“精神主席”，他给自己下达了这项任务。当然，他和他所带的所有炸药一下子就被各种电子的和化学的传感器给盯上了，对他惩罚中的一小部分就是他必须目睹这场开机运行的仪式。

亚当对他的父母都没多少感情，大概这种感情对他来说也没有必要。他所有的13年是在极端奢华的物质中长大的，而这又几乎全部受惠于电脑。他可以得到他所希望的任何东西，只要碰一下按钮即可：食物、饮料、陪伴以及娱乐；而且还有受教育，任何时候只要他感到需要——就会由感人的彩色图像显示来加以说明。他母亲的地位使所有这一切成为可能。

现在，总设计师正在结束他的发言：“……有10^{17} 以上的逻辑单元，这比组成我们国家中任何人的脑部神经的数目还要多！它的智慧将是不可想象的，不过幸运的是我们不必去想象，我们马上就有幸亲眼看到这种智慧：我请我们伟大国家的尊敬的第一夫人，伊莎贝拉·波罗来转动这个开关，让我们的超子电脑开动运行！”

总统夫人向前走去，有点儿紧张，也有点儿笨拙，不过她还是转动了开关。“嘘”的一声，这10^{17} 逻辑单元进入运转时有一丝难以觉察的暗淡的光，每个人都在等待，不知道期望什么是好。“现在有没有观众想提出第一个问题来让我们的超子电脑开始工作？”总设计师问道。每个人都感到羞怯，生怕在众人面前出丑——尤其是在这个新的“上帝”的面前。一片寂静。“可是必须得有一个人来提问呀？”总设计师请求大家。可是大家都害怕，似乎感到了一个新的全权的威慑，亚当可没有这种恐惧。他和电脑一起成长的，他几乎知道作为一台电脑它可能会怎样感觉。至少他自认为他可能知道，不管怎样，他总是好奇。亚当举起手来。“哦，好的，”总设计师说道，“第三排的这位小青年，你要向我们的新朋友提个问题，是吗？”

第1章　电脑能有精神吗

引论

电脑技术在过去的几十年间有了极其巨大的进展。而且，很少人会对未来的几十年内在速度、容量和逻辑设计方面的伟大进步有所怀疑。到那时候，今日的电脑将显得正和我们今天看早年的机械计算机那样的迟钝和初等。其发展的节律几乎是令人恐惧的。电脑已能以人类远远不能企及的速度和准确性实现原先是属于人类思维的独霸领域的大量任务。我们对于机器能在体力方面超过自己早已司空见惯，它并不引起我们的紧张。相反的，我们因为拥有以巨大的速度——至少比最快的田径运动员快4倍——在地球上均匀地推动我们，以一种使几十个人相形见绌的速率挖洞或毁灭废弃建筑的装备而感到由衷的高兴。机器能实现我们过去在体力上从未可能的事，真是令人喜悦：它们可以轻易地把我们举上天空，在几个钟头内把我们放到大洋的彼岸。这些成就毫不伤害我们的自尊心。但是能够进行思维，那是人类的特权。正是思维的能力，使我们超越了我们体力上的限制，并因此使我们比同伙生物取得更加骄傲的成就。如果机器有朝一日会在我们自以为优越的那种重要品质上超过我们，那时我们是否要向自己的创造物双手奉出那唯一的特权呢？

机械仪器究竟能否思维的问题——也许甚至会体会到感情，或具有精神——不是一个什么新问题 [1] 。但是，现代电脑技术时代的来临赋予它新的冲击力甚至迫切感。这一问题触及到哲学的深刻底蕴。什么是思维？什么是感觉？什么是精神？精神真的存在吗？假定这些都存在，思维的功能在何种程度上依赖于和它相关联的身体结构？精神能否完全独立于这种结构？或许它们只不过是（合适种类的）身体结构的功能？无论如何，相关结构的性质必须是生物的（头脑）吗？精神——也能一样好地和电子设备相关联吗？精神服从物理定律吗？物理定律究竟是什么？

这些都被包括在本书我要试图探索的问题之中。要为这么崇高的问题寻求确定的答案当然是无理的要求。我不能提供这个答案：虽然有些人想用他们的猜测强加于我们，但在实际上，任何人也做不到。我自己

的猜测在本书后头将起重要作用，但是我要清楚地把这些猜想和坚实的科学事实区分开来，并且我还要把这些猜想所依据的原因弄清楚。我不如这么说好了，有关在物理定律、数学性质和意识思维的结构之间的关系引起了某些显然全新的问题，我陈述了以前从未有人发表过的观点。这不是我三言两语即能足以描述的观点，因此需要写这么长篇幅的书。但是简略地讲，也许这样会引起一点误会，我至少能说，我的观点认为，正是我们现在对物理基本定律缺乏理解，才使我们不能物理地或逻辑地掌握“精神”的概念。我在这里不是讲，永远不可能很好地掌握这些定律。相反的，本书的部分目的即是企图在这一方面似乎有前途的方向去刺激将来的研究，并且想要提出某些相当特殊的、显然是新的关于“精神”实际上可在我们知道的物理发展中占据什么位置的建议。

我应该清楚地表明，我的观点在物理学家中是非传统的，并因此在目前不太可能被电脑科学家或生理学家所采纳。大部分物理学家会宣称，在人脑尺度下有效的基本定律已经完全知道。当然，在我们物理知识方面一般地仍有许多空白这一点，是无可争议的。例如，我们不知道制约自然的亚原子粒子质量值以及它们相互作用强度的定律。我们还不能使量子理论和爱因斯坦的狭义相对论完全协调——遑论去建立“量子引力论”了。这种理论要使量子理论和他的广义相对论相协调。由于还没有量子引力论，人们就不能理解在已知基本粒子大小的1/100000000000000000000的不可思议的微小尺度下空间的性质，尽管我们以为自己关于比这更大尺度下的知识是足够的。我们也不知道这整个宇宙无论在空间上还是在时间上是有限的还是无限的，尽管这样的不确定性对于人类尺度的物理学似乎没有什么影响。我们不理解既作用于黑洞的核心又作用于宇宙本身大爆炸起源处的物理学。然而，所有这些问题似乎和人类大脑运行有关的“日常”（或稍小一些）尺度问题的距离是要多遥远就有多遥远。它们肯定是遥远的！尽管如此，我将论证，正是在我们鼻子尖（不如说是后面），在我们的物理理解中，正是在和人类思维和意识的运行相关的水平上，还存在巨大的无知！正如我将要解释的，甚至大多数物理学家还不承认这个无知。我还要进一步论断，黑洞和大爆炸与对这些问题的考虑的确有相关之处，这真是令人吃惊！

我将要用证据来支持我提出的观点以说服读者。但是，为了理解这些观点我们还要做许多事。我们将要到奇异的国度以及陌生的研究领域中去旅游。我们要考察量子力学的结构、基础和困惑，狭义和广义相对论、黑洞、大爆炸、热力学定律、电磁现象的麦克斯韦理论以及牛顿力

学的基本特征。当企图要理解意识的性质和功能时，哲学和心理学问题的作用就清楚地呈现出来了。除了设想的电脑模型外，我们当然要对大脑的实际神经生理学稍有些了解。我们要具备人工智能现状的某些观念，还需要知道什么是图灵机，需要理解可计算性、哥德尔定理以及复杂性理论的意义。我们还将深入到数学的基础甚至物理实在的最本质的问题中去。

如果，在这一切的结尾，读者对我要表达的不太传统的论证仍然无动于衷，那么我希望她或他从这个曲折迂回的，但我希望是激动人心的旅途中，得到某些真正有价值的东西。

图灵检验

让我们想象一种新型的电脑被推广到市场上来，它的记忆容量和逻辑单元的数目可能超过了人脑。还假定为此机器仔细地编了程序，并提供了合适种类的大量数据。制造者宣称这种仪器实际上在思4维。他们也许还宣布它们真正是有智慧的。或许他们还走得更远，并提出该仪器实际上感到痛苦、快乐、慈悲、骄傲等，并且自己知道以及实际上理解它们自己的所作所为。的确，它们仿佛就要被宣布是有意识的。

我们如何才能相信制造者的宣称呢？当我们通常买一台机器时，完全根据其所提供的服务来判断其价值。如果它令人满意地完成了我们规定的任务，我们就很高兴。若不是这样，就把它送回去修理或代换。为了检验该制造者所宣称的该仪器实际上具有人类的属性，我们会根据这一判据，简单地判断它在这些方面是否能和人类一样地行为。假定它令人满意地做到这些，我们就没有原因去抱怨制造者，也没有必要把这台电脑退回修理或代换。

这就为我们提供了有关这些事体的非常有效的观点。假定该电脑的动作和一个人在思维时的动作方式不能区分，行为主义者就会说它在思维。我在此刻暂且采纳行为主义者的这一观点。当然，这不意味着我们要求电脑以一个人在思考时可能行走的方式移动。我们更不指望它会活灵活现地像人类：这些和电脑的目的无关。然而，这意味着我们要求它对任何我们介意问它的问题产生拟人的答案。假定它以一种和人类不能相区别方式提供这些答案，则我们就宣称对它的确在思维（或感觉、理

解，等等）这个事实表示满意。

阿伦·图灵在题为《计算机和智力》的著名文章中有力地论证了这一观点。该文于1950年发表在哲学性杂志《精神》上（*Turing 1950*）。（我们以后还要时常提到图灵。）现在称作图灵检验的观念就是首次在这篇文章中描述的。这是为了检验一台机器是否能合情理地被说成在思维的企图。让我们假设一台电脑（正如上面描述的、我们制造者所叫卖的）确实被宣称为在思维。按照图灵试验，该电脑和某个人类的自愿者都躲开到（知觉的）质问者的视线之外。质问者必须依赖向他们双方提出检验问题，来决定两者何为电脑何为人类。这些问题以及更重要的是她 [1] 收到的回答，全部用一种非人格的模式传送，譬如讲打印在键盘上或展现在屏幕上。质问者不允许从任何一方得到除了这种问答之外的信息。人的主体真实地回答问题并试图说服她，他确实是人而另外的主体是一台电脑；但是该电脑已被编好了“说谎”的程序，为了试图说服质问者它反而是人。如果质问者在一系列的这种检验的过程中，不能以任何一致的方式指明真正的人的主体；那么该电脑（或电脑程序，或程序员，或设计者等）肯定是通过了这一检验。

现在人们也许会争辩道，这种检验对于电脑实际上是不甚公平的。因为如果交换一下角色，使人的主体被要求去假装成电脑，而电脑作真实的回答，那么要质问者去发现哪个是哪个就太容易了。她所要做的一切只是要求这些主体进行某些复杂的算术计算。一台好的电脑能够一下子准确地回答，而人很容易被难倒。（然而，人们对此要稍微小心一些。有些“计算奇才”具有非常惊人的心算技巧，从不算错并且显得轻松胜任。例如约翰·马丁·萨查里阿斯·达斯 [2] ，一位文盲农夫的儿子，1824—1861年生活在德国。他能在比1分钟短的时间内用心算完成两个8位数的乘法，或在大约6分钟时间内完成两个20位数的乘法！很容易错认为这是一台电脑在计算。在现代，亚历山大·爱特金和其他人的计算成就也一样地令人印象深刻。爱特金是20世纪50年代爱丁堡大学的数学教授。质问者对此检验所选择的算术问题必须比这个更令人绞尽脑汁，譬如，在2秒钟内乘2个30位数，一台好的现代电脑可轻而易举做到这一点。）

这样，电脑程序员的部分任务是使电脑在某一确定方面比它实际更“愚蠢”。因为如果质问员要问该电脑一个复杂的算术问题，正如我们上面考虑过的，那么现在电脑必须假装回答不了或者马上放弃！但是我

相信以这种方式使电脑变“愚蠢”不是电脑程序员面临的特别严重的问题。使之回答一些最简单的人类不会感到有任何困难的“常识”问题正是他们的主要困难！

然而，在引用这类特例时存在一个固有的问题。因为不管人们会首先提出什么，很容易设想一种方法使电脑正如一个人的样子去回答那个特殊问题。但是，在电脑方面的任何真正理解的缺乏都会因为不断地询问而显露出来，尤其是对于具有创造性和需要真正理解的问题。质问者的一部分技巧在于能设计出如此创造性的问题，另一部分是利用设计来揭示出是否发生某些实在“理解”的探测性的其他问题去追踪它们。她偶尔也可以问一个完全无聊的问题，看看电脑能否检测出差别来，她或者可以加上一两个表面上听起来像是无聊的，而实际上有一点意义的问题，例如她可以说：“我听说，今天上午一头犀牛在一个粉红色的气球中沿着密西西比河飞。你怎么理解此事？”（人们可以想象该电脑的眉头上，泛出冷汗——用一不适当的比喻！）它也许谨慎地回答：“我听起来觉得这不可思议。”到此为止没有毛病。质问者又问：“是吗？我的叔叔试过一回，顺流逆流各一回，它只不过是浅色的并带有斑纹。这有什么不可思议的？”很容易想象，如果电脑没有合适的“理解”就会很快地暴露了自己。在回答第一个问题时，它的存储器可以帮助它想到它们没有翅膀，甚至可以在无意中得到“犀牛不能飞”，或者这样地回答第二个问题“犀牛没有斑纹”。下一回她可以试探真正无意义的问题，譬如把它改变成“在密西西比河下面”，或者“在一个粉红色的气球之中”，或者“穿一件粉红色夜服”，再去看看电脑是否感觉到真正的差别！

让我们暂且撇开是否或何时能造出通过图灵检验的电脑的问题。让我们仅仅为了论证的目的假定，这种电脑已被造出。我们可以问，一台通过检验的电脑是否应该必须说出思维、感觉和理解等。我将要很快地回到这事体上来。此刻我们且考虑它的一些含义。例如，如果制造者的最强的宣布是正确的，就是说他们的仪器是一个思维的、感觉的、敏感的、理解的、意识的生物，那么在我们购买该仪器时就涉及道义的责任。如果制造者的话是可信的，事情就应该是这样子的！开动电脑仅仅是为了满足我们的需要而不考虑其自身的感情应受到谴责。那在道义上和虐待奴隶没有什么差别。一般地说，我们是应避免使电脑经受制造者宣称的它会感觉到的痛苦。当它变得和我们很亲近时，要关掉也许甚至卖掉它，在道义上对我们都是困难的。就会出现因我们和其他人类或其他动物的关系而要把我们卷入的其他无数的问题。所有这些现在都变成

紧密相关的事体。这样，让我们（以及当局！）知道制造者的宣布是否是真的，便具有极大的重要性！我们假定这个宣布是基于他们如下的断言："每一台思维的仪器已被我们的专家严格地进行了图灵检验。"

我仿佛觉得，尽管这些声称的某些含义，尤其是在道义上有明显的荒谬性，但把成功地通过图灵检验当作存在思维、智慧，理解或意识的有效指标的情形，实际上是相当强的要求。如果我们不交谈的话，何以判断他人也具备这些品质呢？实际上还有其他的譬如面部表情、身体运动以及一般动作等判据，它们会大大地影响我们所作的这种判断。但是，我们可以想象（也许在更为遥远的将来）可把一个机器人制造得能成功地模拟所有这些表情和动作。这下子就不必要把机器人和人的主体躲藏在质问者的视界之外，但是质问者随意支配的判据在原则上可和以前相同。

我本人的观点是准备把图灵检验的要求大大地减弱。我似乎觉得要求电脑这么接近地模仿人类，以使得在一种相关的方式下不能和一个人区分开实在是太过分了。我自己想要求的全部是，我们知觉的质问者应该从电脑回答的性质对在这些回答背后的意识存在真正地感到信服，尽管它可能是非常异样的一种意识。这就是迄今建造的所有电脑系统明显缺乏的某种东西。然而，我能觉察到这样的一种危险，如果质问者能决定哪一方事实上是电脑，那么她可能也许是无意识地迟迟不把甚至她能感觉到的意识赋予电脑。或者在另一方面，她也许有这个印象，即她"嗅"到了这个"异物的存在"，即便该电脑没有这种可疑的好处，她还是准备赋予它这个好处。由于这种原因，就在图灵检验原先形式的更大的客观性上，它具有明显的优点，我在下面就一般地遵循于这种形式。我早先提到的对于电脑引起的"不公平"（也就是它必须做人能做的一切才能通过，而人不必会做电脑能做的一切）似乎没有使把图灵检验当作思维等真正检验的支持者忧虑。无论如何，他们的观点时常倾向于不必等太长时间，譬如讲到2010年，一台电脑就能实在地通过这一检验。（图灵原先提出，到2000年，对一位"中等的"质问者仅仅5分钟的提问，电脑的成功率为30%。）这意味着，他们相当有信心，这一不公平不会显著地延迟这一天的到来！

所有这些事体都与根本问题有关：也就是这一操作的观点是否实际上为判断一个对象中存在精神的品质提供一组合理的判据？有些人会竭力争论说它不是。不管模仿得多么有技巧，终究不和实在的东西一样。

我在这一方面的看法是比较中庸。我倾向于相信，作为一般的原则，不管是多么巧妙的模仿，应该总能被足够巧妙的探测检验得出来，尽管这只是信念（或科学乐观主义）而不是已被证明的事实。这样，总的来说，我准备把图灵检验接受为在它的选定范围内是粗略成立的。也就是说，如果电脑对这些问题的确能以一种和人不能区分的方式回答，并如此适当地 [2] 一致地愚弄了我们有理解力的质问员，那么在缺乏任何相反的证据下，我猜想电脑实际上是在思维、感觉等。我在这儿用的这个词，譬如“证据”、“实际上”和“猜想”，其含义是当提到思维、感觉或理解或尤其是意识时，我用这些概念去表明实际客观的“事体”，它在物理形态上的存在与否是我们要确定的某种东西，而不仅仅是语言上的方便！我把这当作一个关键点。我们在所有能得到的证据的基础上作猜测，以辨别这种品质的存在。（这和譬如讲，天文学家想辨别遥远恒星的质量，在原则上没有什么不同。）

必须考虑哪一些反证据呢？关于这一点要预先立下规则是很困难。但是我要弄清楚的是，仅仅说电脑是由晶体管、导线等而不是由神经元、血管等构成的事实本身，我不认为是反证据。我在心里想到的是，在将来的某一时候可以发展出成功的意识理论，这里成功的含义是，它是一个连贯的适当的物理理论，以一种美丽的方式与物理理解的其余部分相协调，而且使它的预言精确地与人类所称何时、是否、到何等程度他们自己觉得是意识相互关联，而且这一理论在考虑我们电脑的想象的意识方面的确关系重大。人们甚至可以摹想按照这一理论的原则建造的“意识探测器”。对于人的主体它是完全可靠的，但在电脑的情形给出与图灵检验相左的结果。在这种情形下，人们必须非常小心地解释图灵检验的结果。我似乎觉得，人们对图灵检验的合适性问题的态度部分地依赖于他对科学技术如何发展的期望。我们以后必须再来考虑其中的一些问题。

人工智能

人工智能是近年来引起人们很大兴趣的一个领域，经常被简写成“AI”。AI的目标是用机器，通常为电子仪器，尽可能地模拟人的精神活动，并且或许在这些方面最终改善并超出人的能力。AI的结果至少在4个方向是有趣的。尤其是有关机器人的研究，它在很大的程度上是为了满足工业对可实行“智力”的，也就是万能和复杂的、原来需要人干预

或控制任务的机械仪器的实际需要，并使他们以超过任何人的能力的速度和可靠性，或者在人类处于危险的各式各样条件下运行。还有专家系统的发展颇具商业和一般的兴趣，在这系统中整个职业的，譬如医学、法律等的主要知识都能编码载入电脑的系统知识库里。这些职业人员的经验和专长能被这种系统知识库所取代吗？所能指望得到的是否只不过是事实的罗列以及意义广泛的前后参考的长长的表格？电脑能否呈现（或模拟）出真正的智慧肯定具有相当大的社会含义。心理学是和人工智能有直接关系的另一领域。人们希望通过利用电子仪器来模拟人脑或某些其他动物头脑的行为，倘若发现做不到，也可以知道有关头脑工作的某些重要的东西。最后，还存在一个乐观的希望，就是由于类似的原因，AI可为精神概念的意义提供洞察，并为其深刻的哲学问题提供一些答案。

迄今为止，AI已进展到何等程度呢？让我来总结是有一些困难。在世界各地存在许多活跃的小组，而我只熟悉这工作的一小部分的细节。尽管如此，可以公正地说，AI虽然的确做了许多聪明的事，但对真正智慧的任何模拟还是非常遥远的事。为了使读者领略这个学科的风味，我首先提到早先的（仍然是令人印象深刻的）成就，以及后来在计算机下棋方面的重要进展。

W.格雷·瓦尔特的“乌龟”[3] 是20世纪50年代早期制造的第一批人工智能仪器之一。在电池快用光前，它会以自己的动力在地面上四处爬行，而且还会爬到离得最近的电插座那儿，把自己插上给电池充电。当充满了电以后，自己会从插座拔出，并重新在地面上爬行！从那时起人们制造出来许多类似的东西（例如，参见Waltz 1982）。另一独立发展的线索是以特雷·维诺格拉德在1972年设计的电脑程序为例子，它在模拟中可以和一堆不同形状和颜色的木块打交道，以不同的顺序和排列把一块放在另一块的上面。尽管早期获得这类成功，但人们证明了，即使设计一个简单的能操纵自如的、避开人类看来“显然”需用动作来躲开的障碍物的、连接的“机器人手臂”也是何等的困难。视场的解释通常也是AI步骤在目前根本无法接近人脑（以及绝大多数其他动物的头脑）似乎能“轻而易举”做到的那种程度的领域。

K. M.柯尔贝在20世纪60年代中期发展的电脑程序也许更和图灵检验相连贯。它如此成功地模拟成一位精神疗法专家，以至于某些病人发现自己宁愿要电脑而不要人类的医疗家，并且对电脑比对人类的对手更

愿把自己的内心袒露得更多些。可将下列（通过电传机）的交谈视作在电脑“医生”和病人之间的这种交流的一个例子[4]：

医生：我可以为你效劳吗

病人：我不知道

医生：请告诉我一些情形

病人：不

医生：请你说得仔细些

病人：今天我厌烦得很

医生：你知道这是什么原因

病人：是的，我老了

医生：请继续讲

病人：仿佛任何事情都激不起我的兴致

医生：请更多地告诉我你的这种感觉

病人：我也略知为何我会有这样的感觉

医生：是吗

病人：我似乎在想有一个家

医生：什么东西使你有这样的想法

病人：它是一种挑战

医生：为什么那对你是重要的

虽然这会给你一种电脑具有某种理解力的可怕印象。在事实上它一

点也没有，只不过是遵循着某种相当简单的机械规则而已。（还存在一个和它相“颠倒”的系统，这时电脑装成一个痴呆症病人，给出教科书的所有答案和症状，并能捉弄一些医科学生，使之相信是一位病人在提供回答！）

计算机下棋也许为机器能显示可认为是“智慧行为”的最好例子。现在（1989年）一些机器在与人对弈时实际已达到极受尊敬的水平，达到了“国际大师”的水准。（为了比较起见，世界冠军卡斯帕洛夫评分高于2700，而这些电脑的评分略低于2300。）特别是，邓恩和卡斯·斯帕拉克连（为忠诚优越商业微处理机）写的电脑程序已达到2110的（Elo）评分，并得到USCF“大师”称号的奖。主要由卡内基·梅隆大学的徐宪编程序的“深思熟虑”更是令人印象深刻，它的评分约为2500Elo，最近在一次下棋锦标赛中（1988年11月，加利福尼亚，长堤）首次取得了（和大师托尼·迈尔斯）共享第一名，并实际上第一回击败了一位大师（本特·拉申）的成就[5]！现在下棋电脑也精于解答棋术问题，它在这方面的造诣轻而易举地超过了人类[6]。

下棋机除了精确的计算能力外，还大大地依赖于“博学多闻”。值得评论的是，只要落子动作要求非常快，下棋机总的来说比相当多的弈手高明一些。如果每一着允许的时间更长，则弈手的表现相对地比机器好。人们可依照如下事实来理解这一切，电脑是基于准确和快速的广义的计算来做决策的，而弈手则依赖于利用相对缓慢的意识评定的“判断”。利用这些人的判断来显著地减少必须在每一计算步骤中认真考虑的可能性，当有时间时，可以得到比不用这类判断而只用简单计算和直接排除可能性的机器更深刻的分析。在玩困难的东方围棋时，这一差别就更显著，那里每一步的可能数目比国际象棋大得多。意识和形成判断之间的关系，将是我后面尤其是第10章论证的中心。

用人工智能得到“快乐”和“痛苦”

人工智能宣称为理解精神品质，譬如快乐、痛苦、饥渴等提供了途径。让我们举格雷·瓦尔特的乌龟为例子。它的行为模式在电池快用完时就要改变，然后它以被设计好的行为方式补充自己的能量存储。这和人类或任何动物感到饥饿时的行为非常类似。当格雷·瓦尔特乌龟以这种方式行为时，说它饥饿了并没十分歪曲语言。其中的某些机制对它电

池的状态很敏感，低到一定点时就会让乌龟转换到不同的行为模式。在动物饥饿时，除了其行为模式的改变更复杂、更微妙之外，无疑存在某些类似的动作。它不是简单地从一种行为模式改变到另一种行为模式，而是存在一种以确定方式行为的倾向的变化，当补充能量供应的需求增加时，这些变化就会更强烈（达到某一点）。

类似地，某些AI的支持者摹想，可以这种方式来适当模拟诸如痛苦或快乐的概念。让我们把情形简化，并只考虑从极端“痛苦”（分数为-100）到极端“快乐”（分数为+100）的单独的“感觉”测度。想象我们有一台仪器，譬如讲是某种电子的、具有记录它自己的（假想的）“快乐——痛苦”度量，我把它称作“苦乐表”。这一仪器具有一定的行为模式和一定的内部的（譬如它的电池状态）或外部的输入。其想法是把它开动以使其苦乐度取最大值。可能会有许多影响苦乐度的因素。我们肯定可以做这样的安排，使得电池中的电荷就是其中的一个因素，低电荷算作负的，而高电荷算作正的，但是还有其他因素。也许我们的仪器装有某些太阳能电池，这是获取能量的另一种手段。这样，当光电池起作用时就不消耗电池的能量。我们可以把光电池朝向光线以增加其苦乐度。这就是不存在其他因素时它所要做的事。（实际上，格雷·瓦尔特乌龟通常避开光线！）我们需要某种实行计算的手段，使得它能弄清它上面部分的不同动作最终在它的苦乐度上的可能效应。可以引进概率权重，使得计算在苦乐度表上具有更大或更小的效应，依其所根据的数据的可靠性而定。

还必须为我们仪器提供仅仅为了维持它的能量供应以外的其他“目的”，否则我们就没有办法去把“痛苦”从“饥饿”中区别出来。在此刻要求我们仪器有生育等功能无疑是太过分了，性的问题不予考虑！但是，我们也许能对它注入一种和其他同类仪器相陪伴的“需求”和它们相遇就得到正的苦乐值。我们或者可以为了其自身的缘故“渴望”学习，使得只要储存有关外部世界的事实即能在苦乐表上得正分（我们可以更自私地安排在为我们作各种服务时得到正分，正如一个人在制造机器仆人时所要做的那样！）。也许有人会争论道，由于凭一时高兴把这种“目的”加到我们的仪器上显得有些做作。但是，这和自然选择加在作为个体的我们身上的，在很大程度上是由于传宗接代的需求所支配的一定“目标”，并没有什么非常大的差别。

现在，假设我们的仪器按照所有这一切已被成功地造出。我们有什

么权利去宣称它的苦乐值为正时它确实感到快乐，而苦乐值为负时感到痛苦呢？AI（或行为主义）的观点是，我们简单地从仪器行为的方式来判断。由于它以一种尽可能增加其正值的（并且尽可能久的）以及相应地尽量避免负值的方式行为，所以我们可以合理地把它的值的正的程度定义为快乐的感觉，而相应地把负值定义为痛苦的感觉。人们会说，此定义的"合理性"正是来自于人类对于快乐和痛苦是以目标方式反应的这一事实。当然，正如我们都知道的，人类的事情实际上并不像这么简单：我们有时似乎特地招惹痛苦，故意回避某种快乐。很清楚，我们的行为实在是由比这些更复杂得多的判据所导引的（参阅*Dennett 1978*，190—229页）。但是作为一个非常粗糙的近似，我们的行为的确是避免痛苦和追求快乐。对于一个行为主义者来说，这已经足够在类似的近似水平上，为我们的仪器的苦乐度和它的痛苦快乐评价的相认同提供正当的理由。这种认同仿佛也是AI理论的一个目的。

我们应该问：在我们的仪器的苦乐度为负或为正时，它是否真正分别地感觉到了痛苦或快乐呢？我们的仪器在根本上是否能感觉到什么呢？行为主义者或者会斩钉截铁地说"显然如此"，或者把这一问题斥为无稽之谈。但是，我觉得这里很清楚地存在一个要考虑的、严肃的困难问题。它对我们自己具有不同种类的影响。有些像痛苦或快乐是可意识的；但是还有其他我们不直接知道的。这可由一个人触摸到热火炉的经验得到清楚的阐明。他在甚至还未感到痛楚之前就采用了抽手回来的不情愿的动作。事情似乎变成，这种不情愿的动作比痛苦或快乐的实际效应更接近于我们仪器对自己的苦乐度的反应。

人们经常用一种拟人化的语言，以一种叙述性的、通常是滑稽的方法来描述机器的行为："今天早晨我的车仿佛不想动"；或"我的手表仍然认为这是加利福尼亚时间"；或"我的电脑宣布，它不理解上一条指令，而且不知道下一步要做什么"。我们当然不是真正地表明车实际上会要什么，或者手表在思维，或者那台电脑[3]真的宣布任何事情，或者它理解甚至知道自己在做什么。尽管如此，假使我们仅仅在它们企图的意义上而不是按字面宣布上接受这样的陈述，则它们可以是描述性的，并且对我们自己的理解有真正的帮助。我将会对AI的有关各式各样建造起来的仪器所具有的精神品质的声称采取类似的态度，而不顾及他们所企图和渴望的！如果我同意说，格雷·瓦尔特乌龟会饥饿，那只是在半开玩笑的意义上这么说的。如果正如上面所摹想的，我准备对苦乐值使用诸如"痛苦"或"快乐"等术语，那只是因为我发现其和我自己的精

神状态的行为有一定的相似性，这些术语有助于我对其行为的理解。我不是暗示这些类似真的是特别接近，或者不存在其他无意识的以更加类似得多的方式影响我行为的东西。

我希望使读者清楚，我的意见是，对精神品质的理解，除了直接从AI得到之外，还存在有更大量的东西。尽管如此，我相信AI实现了一种值得尊敬和慎重处理的严肃的情势。我在说到这些时，并不意味着在人工智能的模拟中，如果有的话，有非常多的成就。但是人们必须心中有数，这个学科还是非常年轻的。电脑的运行会变得更快速，具有更大的可快速存取的空间、更多的逻辑元，并可并行地进行更大数目的运算。在逻辑设计和程序技术方面将会有所改善。这些机器，这种AI哲学的载体将在它们的技术能力方面得到大幅度的改善。此外，该哲学本身也不是固有地荒谬的。也许电脑，也就是当代的电脑的确能非常精确地模拟人类的智慧。这种基于今天被理解的原则，但是具有更伟大得多的能力、速度等的电脑一定会在近年内被制造成功。也许甚至这样的仪器将真正是智慧的；也许它们会思维、感觉以及具有精神。或者它们也许还制造不出来，还需要一些目前完全缺乏的原则。这些都是不能轻易排斥的问题。我将尽我所见地提出证据。我将最终提出自己的看法。

强人工智能和西尔勒中文屋子

有一种称作强人工智能的观点在这些问题上采取相当极端的态度[7]。根据强AI，不仅刚才提到的仪器的确是智慧的并且有精神等，而且任何计算仪器，甚至最简单的机械的，诸如恒温器的逻辑功能都具有某种精神的品质[8]。这种观点认为精神活动只不过是进行某种定义得很好的，经常称作算法的运算。下面我将精确地说明算法实际上是什么。此刻暂且把算法简单地定义为某种计算步骤就已足够了。在恒温器的情形下，其算法至为简单：仪器记录其温度是否比设定的更高或更低，然后使线路在前面情形时断开，而在后面情形时接通。对于人脑的任何有意义的精神活动，其算法必须比这复杂得多，它和恒温器的简单算法在程度上具有极大的差别，而在原则上则相同。这样，根据强AI，人脑的主要功能（包括它的一切的意识呈现）和恒温器之间的差别只在于，在头脑的情形中具有大得多的复杂性（或许“更高级的结构”，或“自省性质”，或其他可赋予算法的属性）。按照这一观点，至为重要的是，所有精神品质，譬如思维、感情、智慧、理解、意识都仅仅被认为是这一

复杂功能的不同侧面；也就是说，它们仅仅是头脑执行的算法的特征。

任何特殊算法的价值在于它的表现，也就是它的结果的精确，它的范围，它的经济性和它可运行的速度。一种想和人脑中假想的运行的算法相比拟的算法一定是非常了不起的东西。如果头脑中存在有这一类算法，强AI支持者肯定作此断言，那么在原则上它可在一台电脑上执行。假定它不受存储容量和运算速度的限制的话，的确可在任何当代的通用电脑上执行。（我们以后去考虑普适图灵机时，这一评论就会得到证实。）人们预料，在不太远的将来大型快速电脑将会克服任何这类限制。一旦这样的一种算法能被找到，它就能通过图灵检验。强AI支持者就会宣布，只要执行该算法，它自身就会体验到感情，具有意识，并且是一种精神。

绝不是每一个人都同意，可用这类方法把精神状态和算法相等同。美国哲学家约翰·西尔勒（1980，1987）尤其反对这种观点。他引用过这种例子，即假定有一台适当地编了程序并已经实际上通过了简化的图灵检验的电脑；但是他以有力的论证支持如下观点：即便如此，这台电脑仍然完全不具备和理解有关的精神属性。其中一个例子是基于罗杰·施安克（*Schank and Abelson 1977* ）设计的电脑程序之上。该程序的目的是为理解简单的故事提供模拟，例如：“一个人进入餐馆并订了一份汉堡包。当汉堡包端来时发现被烘脆了，此人暴怒地离开餐馆，没有付账或留下小费。”第二个例子是：“一个人进入餐馆并订了一份汉堡包。当汉堡包端来后他非常喜欢它；而且在离开餐馆付账之前，给了女服务生很多小费。”作为对“理解”这一故事的检验，可以询问电脑，在每一种情形下此人是否吃了汉堡包（这一事实在任一故事中都没有说清）。电脑对这类简单的故事和问题可给出和任何讲英文的人都会给出的根本无从区别的回答，也就是对于这些特定的例子，第一种情形是“非”，而第二种情形是“是”。这样一台机器已在这一非常有限的意义上通过了图灵检验！

我们应该考虑的问题是这类成功是否实际上表明电脑方面或许程序本身方面具有任何真正的理解。西尔勒使用了他的“中文屋子”的概念来论证它不具备。他首先摹想，这一故事是用中文而不是英文来讲，这肯定是非本质的改变。把这一特殊演习的电脑算法的所有运算（用英文）作为一组指令提供给用中文符号进行操作的计算员。西尔勒想象自己被锁在一个屋子里操纵这一切。代表这一故事和问题的一连串符号通过一

条很小的缝隙被送进这屋子。不允许任何其他的来自外面的信息漏进去。最后当所有的操作完成后，程序的结果又通过这条缝隙递到外面来。由于所有这些操作都是简单地执行施安克程序的算法，这个最终程序的结果简单地为中文的“是”或者“非”，给出了关于以中文说的故事用中文问的原先问题的正确答案。现在，西尔勒很清楚地表明他根本不识中文，这样他对该故事讲的是什么没有任何哪怕是最浅显的概念。尽管如此，只要正确地执行了那些构成施安克算法的一系列运算（已给他用英文写的这一算法的指令），他就能和一位真正理解这故事的中国人做得一样好。西尔勒的要点是，而且我以为是相当有力的，仅仅成功执行算法本身并不意味着对所发生的有丝毫理解。锁在他的中文屋子里的（想象的）西尔勒不理解任一故事的任一个词！

人们对西尔勒的论证提出了许多异议。我将只提到我认为具有重要意义的那些。首先，在上面用到的“不理解任一个词”的短语也许有容易使人误导的东西。理解和模式之间正与它和单独词汇之间有一样多的关系。在执行这类算法时，在不理解许多个别词汇的实在意义的情形下，人们可以知觉这些符号构成的模式的某些东西。例如，中国字的“汉堡包”（如果真的有这个词的话）可用某一其他的食品譬如讲“炒面”来替换，而故事不会受到重大影响。尽管如此，我觉得可以合理地假设，如果人们仅仅跟踪着这种算法细节的话（即使认为这种代换不重要），则只传递了该故事中很少的实际意思。

其次，人们必须考虑这个事实，如果用人类的操纵符号来执行的话，在正常情形下甚至执行一个相当简单的电脑程序也会是非同寻常的冗长和繁琐。（这毕竟正是我们让电脑来为人类做这种事的原因！）如果西尔勒真的以这种提议的方式实行施安克的算法，那么仅仅是为了得到哪怕是一个单独问题的答案，他很可能要花费许多天、许多月甚至许多年极其枯燥的工作，这根本不像是一位哲学家的活动！然而，由于我们在这里主要关心原则的而不是实践的事体，所以我仿佛觉得这不是一个严重的反对。在具有和人脑相当的足够的复杂性、并因此适当地通过图灵检验的假想的电脑程序中产生了更多的困难。任何这类程序都是极可怕的复杂。人们可以想象，为了回答甚至相当简单的图灵检验问题，这一段程序的运算会涉及如此多的步骤，以至于在一个正常人的一生中根本没有可能用手完成这一算法。在没有这种程序的情形下，这是否的确如此还很难说 [9] 。但是，依我的观点，这一极其复杂的问题无论如何不能简单地不予理睬。是的，我们在这里关心的是原则的事体，但是

一个算法要呈现出精神品质，其复杂性就要达到某一“临界”量，我认为这是合情理的。这一临界量也许是如此之大，复杂到这等程度的算法，由任何人以西尔勒摹想的样子用手来进行，根本就是不可想象的。

西尔勒本人允许一整队不能讲中文的符号操作员去取代原先中文屋子里的孤独者（“他自己”），以此来抵抗上面的反对。为了得到足够大的数目，他甚至想象把它的屋子用印度整个国家来代换，现在它的全部人口（除了理解中文的人以外！）都来从事符号操作。虽然这在实践上是荒唐的，但在原则上却不是，而且该论断在本质上和以前是一样的：尽管强AI宣称只要实现适当的算法即会诱导出“理解”的精神品质，这些符号操作员仍然不理解这故事。然而，现在另有一种反对正开始幽然逼近。这些单独的印度人难道不是比起来更像人脑中的神经元而不像整个人脑本身吗？没人认为神经元本身会单独地理解这个人的思想，神经元的激发明显地构成人脑在进行思考时的物质活动，为何要期望单个印度人去理解这中文的故事呢？为了答复这一诘问，西尔勒指出，如果在印度没有一个单独的人理解这一故事，而真实的国家却能理解显然是荒唐的。他论证道，一个国家正像一台恒温器或一辆汽车与“理解”毫不搭界，而与单独的个人却有关系。

这一论证，比前面的那个要苍白无力得多。我认为，西尔勒的论证在只有一个单独的人在实行算法时力量最强大，这时我们只限于注意一个不复杂到可由一个人在短于一生的时间内实际执行的算法的情形。我认为他建立这一结论的论证不够严格，这就是不存在和一个人实行那个算法相关的离体的某种类型的“理解”，而且这种理解的存在并不以任何方法反射到他自身的意识上去。然而，我和西尔勒都同意，至少可以说，这种可能性被减少到很微小的程度。我认为西尔勒的论证对之还有相当的力量，即使它还不完全是结论性的。施安克的电脑程序所具有的这类复杂性的算法不能对其实行的任何任务有丝毫真正的理解，对这一点的展示是相当令人信服的，而且它（仅仅）暗示，不管一种算法是多么复杂它都不能自身体现真正的理解。这和强AI的声称相矛盾。

就我所能看到的，强AI观点中还有其他非常严重的困难。强AI观点就只管算法。这个算法是由头脑、电脑、印度整个国家、轮子和齿轮还是由一套水管系统来执行都是一样的。其观点是，对于被认为由算法所代表的“精神状态”，只有它的逻辑结构是有意义的，这与那个算法的特殊的物理体现完全无关。正如西尔勒所指出的，这在实际上导致了一

种“二元论”的形式。二元论是由极富影响力的17世纪的哲学家兼数学家雷恩·笛卡儿所提倡的，它断言存在物质的2种不同的形式：“精神的东西”和通常物质。这两类物质的一种是否并且如何去影响另一种是个额外的问题。关键是认为精神的东西不是由物体所构成，并能独立于它而存在。强AI的精神东西是算法的逻辑结构。正如我刚评论过的，一个算法的特殊的物理体现，是完全无关的某种东西。算法有某种离体的“存在”，这和它的按照物理的实现完全分离。我们要多么认真地对待这种存在是我在下一章还要讨论的一个问题。它是抽象数学对象的柏拉图实在的一般问题的一部分。此刻我且回避这一般的问题，而且仅仅评论强AI支持者所仿佛的确相信的算法形成它们思维、感情、理解以及意识、知觉的“物质”。正如西尔勒指出过的，强AI的立场似乎把人们逼向极端的二元论，也就是强AI支持者最不愿意与之打交道的观点，这真是富有讽刺意味！

这就是在道格拉斯·霍弗斯达特（1981）的题为《和爱因斯坦头脑谈话录》的对话中所论证的事件背后的两端论。他本人是强AI的主要提议者。霍弗斯达特捏造出一本极厚的书，假想它能包含对阿尔伯特·爱因斯坦头脑的整个描述。正如活着的爱因斯坦要回答的那样，只要简单地翻阅该书，并仔细地按照所提供的细致的说明，任何人愿意问爱因斯坦的问题都能得到回答。当然，正如霍弗斯达特所指出的那样，“简单”是十足的误称。但是他的宣称是，在图灵检验的操作意义上，这本书在原则上完全等效于实在的爱因斯坦的慢得可笑的复件。这样，按照强AI的论争，这本书可像爱因斯坦本人那样思维、感觉、理解和知觉，但是也许是以极慢的节律生活（这样，从书——爱因斯坦看来，外部世界似乎以疯狂的高速度闪现）。由于这本书被认为仅仅是组成爱因斯坦“自己”的算法的特殊体现，它实际上就是爱因斯坦。

但是，现在出现了新的困难。这本书也许从未被打开过，或者它被无数追求真理的学生和研究者熟读。这本书怎么“知道”这种差别呢？这本书也许不必被打开，它的信息可由X射线立体扫描术或其他技术的魔法取出。爱因斯坦的知悉是否只有当这本书被考察时才被唤起呢？如果两个人在两个完全不同的时刻去问该书同样的问题，他是否发觉是发生了两回？或者那是否使爱因斯坦知觉的同一状态在两个分开的和时间不同的事件中实现？或许只当该书被改变时，他的知觉才被唤起？毕竟在正常情况下，当我们知觉到从影响我们记忆的外部世界接受到一些信息时，我们的精神状态确实稍被改变。如果是这样的话，这是否意味着，

算法的"适当的"改变（我在这里把记忆的储存包含到算法部分中去）而不是（或者以及）算法的激活被认为和精神事件相关？或者即便书——爱因斯坦从未被任何人或任何东西考察或扰动过，它是否维持在完全自知觉的状态？霍弗斯达特触及了其中的一些问题，但他完全不想去回答这些，或者和我们大多数人妥协。

去激活或者以物理形式体现一个算法是什么意思呢？改变一个算法是否和仅仅抛弃一个算法并且用另一个取代之在任何意义上是否不同呢？这些究竟和我们意识知悉的感觉有关吗？读者（除非他或她本人是强AI支持者）也许会惊讶，为什么我为了这样明显的荒谬的思想花了这么多的篇幅。事实上，我不认为这一思想在本质上是荒谬的，它仅仅是错误的！在强AI的背后的推理中的确有某种必须慎重对付的力量，我将解释这一点。我的观点是，这些思想若被适当地修正的话，还具有一些魅力，正如我还将告知诸位的。此外，依我的看法，西尔勒表达的特殊的矛盾的观点还包含某些严肃的困惑和表面的荒诞，尽管如此，在一定程度上我仍然同意他！

西尔勒在他的讨论中似乎隐含地接受当代的电脑，但是还加上大大加快了的动作速度和快速存取的记忆容量（而且可能并联运行），可以在不太遥远的将来体面地通过图灵检验。他是准备接受强AI的论点（以及其他大部分"科学"观点），"我们是任何电脑程序的体现"。此外，他还附和这种说法："头脑理所当然地是一台电脑。由于任何东西都是一台数字电脑，头脑也是[10]。"西尔勒坚持，人脑（它可有精神）和电脑（他论证说没有精神），两者都可以执行同样算法，两者功能之间的差别完全在于各自的物质构成。但是，由于他不能解释的一种原因，他声称生物体（头脑）可有"意图性"和"语义性"，他把这些定义为精神活动的特征，而电子仪器没有。我觉得这对于得到科学的精神理论实质上没有什么用处。也许除了生物系统（而我们刚好是这样的系统）的进化来的历史的"方式"以外，关于它有什么特殊的东西特地被恩准获得意图性或语义性？我觉得这一断语就像教义一样地令人可疑，甚至也许不比强AI的坚持的只要执行一个算法即能召唤起意识知觉的状态的说教更加独断！

我的意思是，西尔勒以及大量其他的人被电脑专家引入歧途。而这些电脑专家又依序地被物理学家引入歧途。（这不是物理学家的过错。他们甚至不知道发生了什么事情！）"每一件东西都是一台电脑"的信念

似乎已广泛蔓延。我在本书的愿望是为了表明，为何以及如何情况并非如此。

硬件和软件

在电脑科学的行话中，术语硬件表示一台电脑，涉及的实在的机构（印刷线路、晶体管、导线、磁性存储空间等），包括了把所有东西都联结起来的方法的全部细节。相应地，术语软件是指可在机器上进行的各种程序。阿伦·图灵的一个杰出的发现便是，任何其硬件达到一定程度复杂性和灵活性的机器，都等效于任何其他同类机器。这一等效性的含义是，对于任何两台这样的机器A和B，存在一段特别的软件，如果将其赋予机器A，就会使之完全像机器B一样地动作；类似地，还存在另一段特殊的软件，如果将其赋予机器B，就会使之和机器A完全一样地动作。我在这里用词“完全一样”是指，对任何给定的输入（在对机器提供了转换软件之后再提供它）机器的实际输出，而不是指每台机器用以产生这输出所花的时间。如果任何一台机器在任何阶段用光了用于计算的存储空间，我还允许其调用一些（在原则上无限制的）外部空白空间的“粗纸”供应，可采用磁带、磁盘、磁鼓或任何别的什么。事实上，对机器A和B执行同一任务所需时间的不同值得严肃地加以考虑。例如，也许有这种情形，A在执行一特别任务时比B快1000倍；也许还有这种情形，同样的一对机器，存在某一其他任务，这时B比A快1000倍。此外，这里计时可以极大地依赖于所用的转换软件的选取。这是非常“原则的”讨论，人们不甚关心诸如在一段合理的时间内完成他的计算的实际的事体。我将在下一章把在这里提到的概念弄得更精确些：机器A和B是所谓通用图灵机的实例。

实际上，所有现代通用的电脑都是通用图灵机。这样，在上述的意义上，所有通用的电脑都是互相等效的：假使我们不关心运算的速度和存储空间的可能限制，则它们之间的差别可完全地被包摄到软件中去。的确，现代技术已经使得电脑如此快速地运行，并具有如此庞大的储存能力，对于大多数“日常”目的，这些实际考虑对于通常的需求不构成任何严重的限制 [4] ，所以这种电脑之间有效的理论上的等价也可认为是在实践的水平上的。技术仿佛已经把有关理想化的计算仪器的纯学术讨论转变或直接影响我们生活的事体。

就我所知，作为强AI哲学基础的最重要因素之一是在物理计算仪器之间的这种等价。硬件似乎相对地不重要（也许甚至完全不重要），而软件也就是程序或者算法被认为是要紧的因素。然而，我似乎觉得还有从物理的方向来的更多的重要的基础因素。我将指出这些因素是什么。

是什么东西赋予个别人其单独的认同性呢？在一定的程度上，是否正是构成他身体的原子呢？他的认同性是否依赖于构成这些原子的电子、质子和其他粒子的特殊选择呢？至少有两种理由说明不可能是这样子的。首先，任何活人身体的物质都处于连续代换的状态中。这尤其适用于一个人脑的细胞中，尽管在出生后没有产生新的脑细胞的这一事实。在每一活细胞（包括每一个脑细胞）中的绝大多数原子以及实际上我们身体的整个物质从诞生以来已被代换了许多回。

第二个理由来自于量子物理。而且极富讽刺意味的是，严格地讲，它和第一个理由相冲突！按照量子力学（我们在第6章还要进一步地讨论），任意两个电子必须是完全等同的，这同样地适合于任意两个质子以及任一特殊种类的两个粒子。这不仅仅说没有办法把两个粒子区分开，其陈述比这还要强许多。如果一个人脑中的一个电子和一块砖头中的一个电子相互交换，则系统的态和它过去的态不仅不能区分，而且完全相同 [11] 。这同样适用于质子和任何其他种类的粒子，整个原子、分子等。如果一个人的整个物质内容和他房子里的砖头的相应的粒子相交换，那么在某种强的意义上来讲，没有发生过任何事情。把人和他的房子区分开来的是把这些成分安置的模式，而不是这些成分本身的个性。

在与量子力学无关的日常水平上也许存在一个相似的情景。这就是由于电子技术使得我能以文字处理机来打字。当我写到这里时感到特别明显。如果我必须改变一个词，譬如说把“make”改成“made”，我只要简单地把“k”用“d”来取代即可以，或者我可重打整个词。如果我重打，新的“m”是否和旧的“m”一样，或者我是否用同样的字母来取代了它呢？“e”的情形又如何呢？即使如果我简单地用“d”来取代“k”，而不重打这个词，存在刚好在“k”消失和“d”出现从而填上空隙之间的一个瞬息，随着接续的每一字母（包括“e”）的安置，存在（或者至少有时存在）重排这页之后的波动。然后当“d”插进去时又再次重新计算。（呵，现代没有思想的计算是多么的“卑贱”！）不管怎么样，在我面前屏幕上看到的所有字母，随着它的每1秒钟的60次扫描，仅仅是一个电子束的轨迹的缝隙。如果我取走了任一字母并用同一字母取代之，在代换后的情

形是否一样，或仅仅和原先的不可区分？认为第二种观点（也就是“仅仅是不可区分的”）可以和第一种观点（也就是“同样”）相区分似乎是痴人说梦。至少在字母不变时可以合理地说这情形是同样的。而等同粒子的量子力学的情形也是如此。把一个粒子用另一个等同粒子取代时量子态不受丝毫影响。这情形的确被认为和以前的是同样的。（然而，正如我们在第6章将要看到的，这一差异在量子力学的框架中实际上不是微不足道的。）

上述关于在一个人体中连续地置换原子的评论，是在经典的而不是量子物理的框架下进行的。它是在似乎坚持每一原子的个性有意义的情形下措辞的。经典物理在这一描述的水平上把原子当作单独物体的近似是足够好的，我们不会错得太离谱。假设原子在运动时和它们等同的伙伴分离得相当开，那么由于在事实上每一个原子的轨道是连续的，以致人们想象能够看守住每一个并可以协调地认为它们坚持各自的本体。从量子力学的观点看，原子的个性只不过是一种方便的说法，但是在刚才考虑的水平上它是一个足够协调的描述。

让我们接受这种观点，一个人的个性和人们想赋予他的物质成分的任何个性无关。相反地，在某种意义上，它必须和那些成分的形态，即我们所说的空间或空间——时间中的形态有关。（后面还要更多地讲到。）但是强AI的支持者走得比这更远。如果这样一种形态的信息内容能被翻译成另一种可能恢复成原状的形式，那么他们就能够宣称，这个人的个性必须维持不动。这正如同我刚打的字母序列和我现在于我的文字处理机屏幕上看到的展示是一样的。如果我把它们从屏幕上移开，它们被编码成某种微小的电荷位移的形式，处在一种和我刚才打印的字母在几何上毫无类似性的某种形态。然而，我可在任一时刻把它们移回到屏幕上去，它们在那里正如同没有进行过任何变换一样。如果我选择把我才写下的存起来，那么我可以把字母序列的信息转移到一个以后我可取走的磁盘的磁化形态上去，然后关掉机器就中和了在它上面的所有（有关的）微小电荷位移。第二天，我可重新插入我的磁盘，复原小电荷位移并在屏幕上正如没有发生过任何事一样重现字母序列。这对于强AI支持者而言是“清楚的”，一个人的个性可用同样的方式处理。所以这些人会宣称，正如在我的显示屏幕上的字母序列一样，如果一个人的身体形状被翻译成完全不同的某种东西，譬如说一块磁铁的磁场，他的个性一点也没损失，实际上对他根本没有发生过什么。他们甚至会宣称，当一个人的“信息”处于这不同的形式时，他的意识知觉继续存留。一

个“人的知觉”在这种观点中实际被当成一段软件，而他作为一个物质的人的特殊的显现则被认为是通过他头脑和身体硬件对这软件的运算。

作这些断言的原因仿佛是，不管硬件采取何种物质形式，例如某种电子仪器，人们总可以“问”软件问题（以图灵检验的形式），并假定该硬件能令人满意地进行计算以获得这些问题的答案。这些答案会和一个人处于正常状态时所回答的相同。（“今天上午你感觉如何？”“哦，相当好，谢谢，尽管我有一点讨厌的头痛。”“你对你的个人认同感和别的什么有点不对头的地方吧？”“不，你为什么这样讲？这似乎是很古怪的问题。”“那么你感到你正是昨天的那个同一的你吧？”“当然是这样！”）

科学幻想的超距运送机 [12] 是一种被频繁讨论的观念。这是作为譬如讲从一颗行星到另一颗行星的“运送”手段。所有的讨论都是关心它是否能在实际上做到这样。旅行者不用空间飞船以“正常”方式运送其身体，而是从头到脚地被扫描，他身体的每一原子和电子的准确位置和完整的特征都被全部细致地记录下来。然后所有这些信息由一电磁信号束（以光速）发射到目的地。在目的地把这信息收集到，并作为装配旅行者以及他所有记忆、企图、希望和最内心的感情的复本的指导书。至少这就是所期望的，因为他的头脑状态的每一细节都被完全忠实地记录、传送和重造了。假如这个机制能成功，则旅行者的原版可被“安全地”毁掉。显然的问题是：这真的是从一处到另一处旅行的一种方法吗？或者它是否仅仅是制造一个复本而把原先的人杀死？假定这种方法在这框架中被证明是完全可靠的，你会准备用这种方法吗？如果超距运送不是旅行的话，那么在原则上它和从一个房间走到另一个房间有何不同？在后者的情形，难道不是一个时刻的原子简单地为下一时刻的原子提供定位的信息吗？我们毕竟看到了，维持任何特殊原子的等同性都是没有任何意义的。原子的任何运动模式难道不就是构成从一处到另一处传播的信息的波吗？在描述从一个房间随便遛达到另一房间的我们旅行者的波动传播和发生在超距运送机中的有什么本质的区别吗？

假定远距运送的确“可行”，那就是说，在一遥远行星的旅行者的复本中他自身的“知觉”确实被重新唤醒（假设这是一个具有真正意义的问题）。如果该旅行者的原版没有如这个游戏所需求的那样被毁坏，将会发生什么呢？他的“知觉”是否会同时处于两个地方呢？（当你听到如下一段话时想象着将会如何反应：“哦，亲爱的，在把你放到超距运送机

前给你的药已被消耗完了。是吗？那是有点不幸，但是没有关系。无论如何，你将高兴地听到，另一个你，呃，我是说真正的你已经安全地到达了金星。这样，我们可以把你，哦，我指的是多余的版本，安置在这儿。当然这是完全不痛苦的。”）这种情景有一点佯谬的风味。物理学的定律中是否有任何在原则上使得超距运送不可能的东西呢？另一方面，也许在原则上没有东西反对以这种手段运送一个人以及他的意识，但是所涉及的“复制”过程会不可避免地消灭原来的那个人吗？尽管这些考虑显得很奇异，我相信从它们也许可得到某些关于意识和个性的物理性质的东西。我相信它们提供了表示量子力学在理解精神现象的某种根本作用的指针。但是我更往前多跨一步。我们只有在第6章（参见第325页）考察了量子理论的结构后，才能回到这些事体上来。

让我们看看强AI的观点与远距运送问题有什么关联。我们设想，在这两个行星之间的某处有一转换站，在这里把信息暂时存储然后再传送到最终的目的地。为了方便起见，这信息不用人的而是用某种磁或电的仪器的方式存储。该旅行者的“知觉”是否会在和这一仪器的相关联中呈现呢？强AI的支持者愿使我们相信，事情必须如此。他们说，我们想问该旅行者的任何问题，在原则上都可由此仪器答复，“只要”对他的头脑的适当活动建立模拟就可以了。该仪器会拥有所有必需的信息，而余下的只不过是计算的问题。由于仪器会完全如同旅行者一样地回答问题，那么（图灵检验！）它就是该旅行者。这完全回到了强AI的论点，在考虑精神现象时硬件根本不重要。我觉得这一论点是未被证实的。它是基于如下的假设，即头脑（或精神）实际上是一台数字电脑。他们假想，当一个人思维时并没有引起特别的物理现象，也许头脑真正需要具备特殊的物理的（生物的、化学的）结构。

人们无疑地会（从强AI观点）争论道，所做的仅有的假设是，任何必须涉及的特殊物理现象的效应都可由数字电脑精密地仿照。我可以相当肯定，大多数物理学家会论证道，在我们现在对物理理解的基础上作这样的假设是非常自然的。我将在后面的章节提出我自己持相反观点的原因（我在那里还需要把话题引到为何我相信甚至不必作任何假定）。但是，在此刻我们暂且接受这一（普遍的）观点，即所有相关的物理总能由数字计算来仿照。那么（除了时间和计算空间的问题外）这唯一真正的假设是一个“行为主义”的问题，即如果某物全然像一个意识的知觉的本体那样地行为，那么人们还应该坚持说它“感觉”到它自己是那一个本体。

强AI观点认为，在头脑运行中实际上被涉及的任何“仅仅”是作为硬件问题的物理，必须能用合适的转换软件来模拟。如果我们接受这个行为主义的观点，那么问题就归到普适图灵机的等价，以及任何算法可的确由这种机器执行的事实，还有头脑按照某类算法动作行动的假设。现在到了我要更明白地解释这些迷人的重要概念的时候了。

第**2**章　算法和图灵机

算法概念的背景

算法、图灵机或者普适图灵机究竟是什么呢？为何这些概念在可以构成“思维仪器”的东西的现代观点中占有如此核心的地位呢？是否在原则上存在一个算法可达到的绝对极限呢？为了充分地讨论这些问题，我们必须比较细致地考察算法和图灵机的观念。

我在下面的各种讨论中，有时将要用到一些数学表达式。我注意到有些读者排斥这类东西，或者觉得它们吓人。如果你是这种读者，那么我请你原谅，并请你参照我在“敬启读者”中的建议。其实，这里论证所需的数学知识并不超过小学水平，但要仔细地弄通它们，则需要一些认真的思考。事实上，大部分描述是十分显明的，只要细心地跟随就能很好地理解。即使，如果人们只是为了稍微领略其风味而取其精华，也能有很大的收益。另一方面，如果你是一位专家，我还要请你原谅。我猜想，它仍然值得你花一段时间把我所写的看一遍，并且可能会有一两件东西引起你的兴趣。

“算法”这个词来自于9世纪波斯数学家阿布·雅发·穆罕默德·依伯恩·缪莎·阿勒——霍瓦里松，他在公元825年左右写了一本影响深远的《代数对话录》。“算法”这个词现在之所以被拼写成“algorithm”，而不是早先的更精确的“algorism”，似乎是由于和“算术”（arithmetic）相关联的缘故。[还值得指出的是，“代数”（algebra）这个词来源于该书书名的阿拉伯字“al jabr”。]

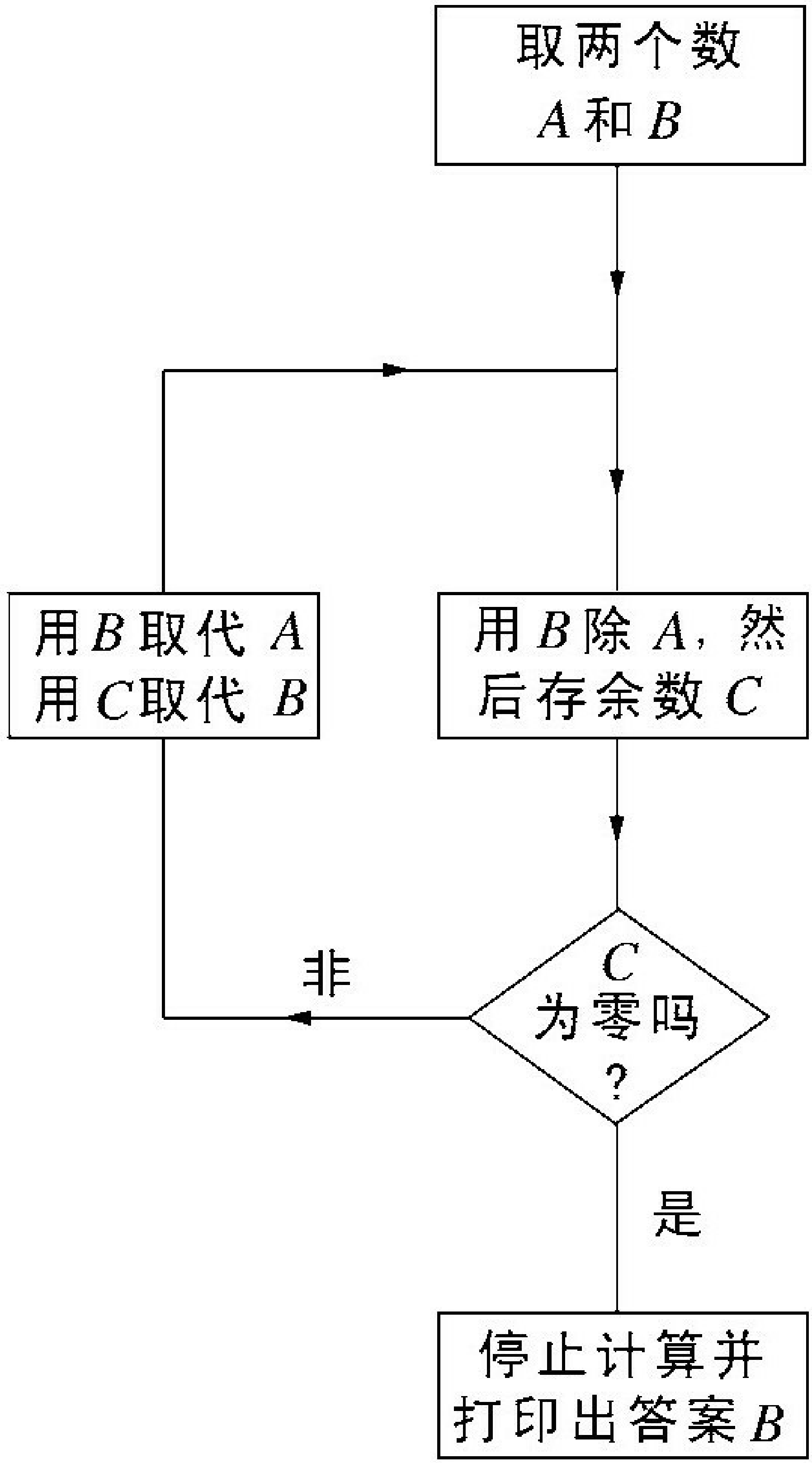
然而，在阿勒——霍瓦里松的书以前很久人们就知道了算法的实例。现在被称作欧几里得算法的找两个数最大公约数的步骤是在古希腊（公元前300年左右）即有记载的一个非常熟知的例子。让我们看看这是如何进行的。随意取两个具体的数，譬如讲1365和3654。所谓的最大公约数是可以同时整除这两个数的最大的整数。在应用欧几里得算法时，我们让这两数中的一个被另一个除并取余数，在3654中取出1365的两倍，其余数为924（=3654-2730）。我们现在用此余数即924以及我们

刚用的除数即1365去取代原先的两个数。我们用这一对新的数重复上述步骤，用924去除1365，余数为441。这又得到新的一对441和924，我们用441除924，得到余数42（=924-882），等等，直到能够被整除为止。我们把这一切如下列出：

$3654 \div 1365$	得到余数924
$1365 \div 924$	得到余数441
$924 \div 441$	得到余数42
$441 \div 42$	得到余数21
$42 \div 21$	得到余数0

我们最后用于做除数的21即是所需要的最大公约数。

欧几里得算法本身是我们寻找这一因子的系统步骤。我们刚才把这一步骤应用于具体的一对数，但是这步骤本身可被十分广泛地应用于任意大小的数。对于非常大的数，要花很长时间来执行该步骤，数字越大则所花的时间越长。但在任何特定的情形下，该步骤最后会终结，并在有限的步骤内得到一个确定的答复。每一步骤所要进行的运算都是非常明了的。此外，尽管它可应用到大小没有限制的自然数上去，但是可以用有限的术语来描述整个过程。（“自然数”简单地就是通常的非负[1]整数0，1，2，3，4，5，6，7，8，9，10，11……）。的确很容易建立一个（有限的）描述欧几里得算法全部逻辑运算的“流程图”。



取两个数
A和B
用B取代 A
用C取代 B
用B除 A，然
后存余数 C
C
为零吗
?
非
是
停止计算并
打印出答案B

应该提到，我们在这里隐含地假定，已经“知道”如何实行从两个任意自然数A和B的除法中得到余数的必需的基本运算，所以这一步骤还未完全被分解成最基本的部分。那种运算又是算法的，是用我们在小学学到的除法的非常熟悉的步骤来进行的。在实际上，这个步骤比欧几里得的其他部分更复杂不少，但是可以再为它建立一个流程图。其复杂性主要起源于这个事实，即我们是（假定）对自然数用标准的“十进制”记数法。这样子我们需要列出全部的乘法表并考虑进位，等等。如果我们简单地用一连串n个某种记号来代表数n，例如×××××代表5，那么可从非常初等的算法运算看到余数的形成。为了得到当A被B除时的余数，可以简单地从代表A的记号中不断去掉代表B的符号串，直到最后余下的记号不够再进行这种运算为止。最后剩下的符号串提供了所需的答案。例如，为了得到17被5除的余数，我们可以简单地从×××××××××××××××××不断地取走5的序列×××××正如下面所示：

×××××××××××××××××

××××××××××××

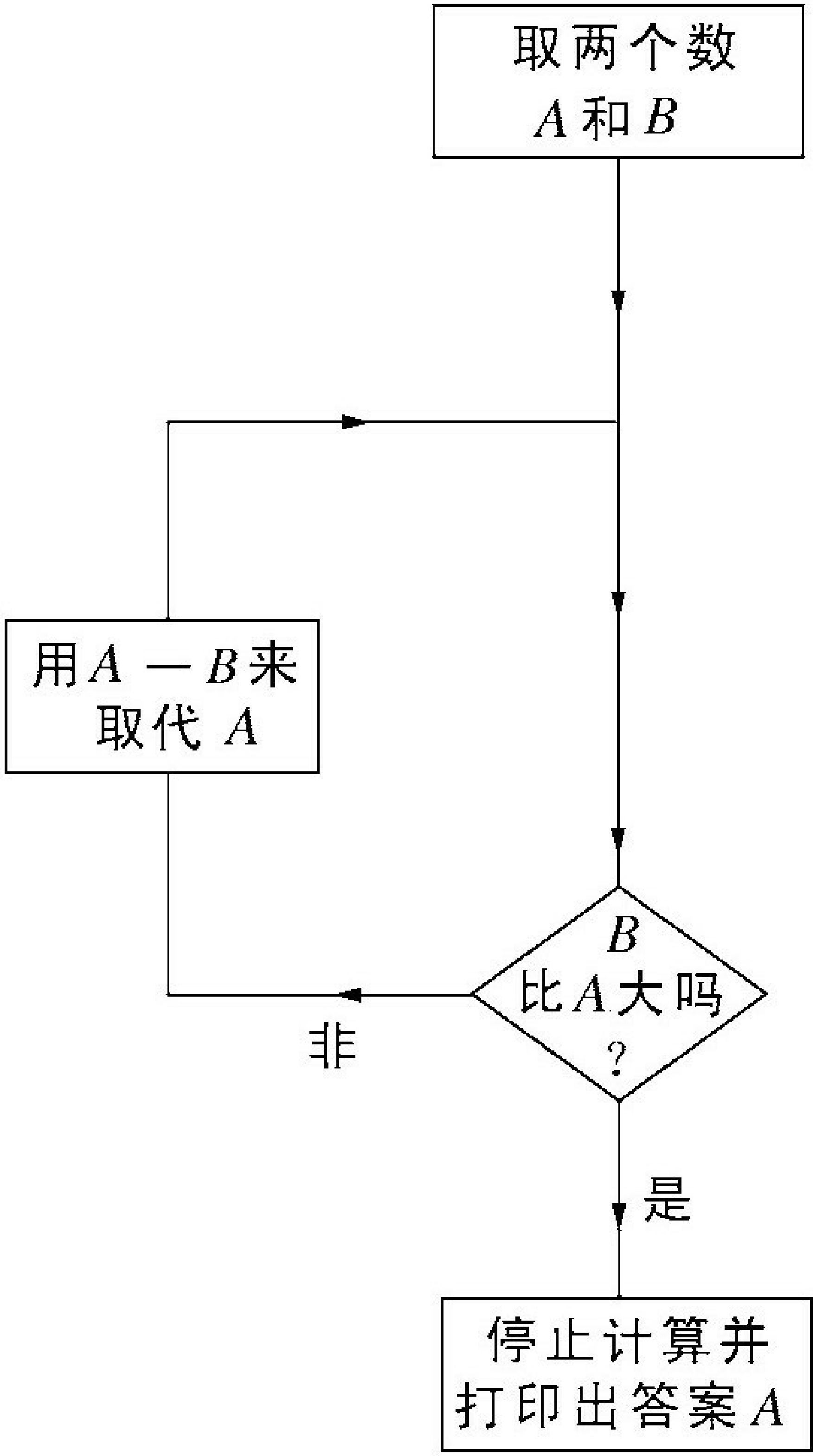
×××××××

××

由于我们不能再继续这种运算，所以很清楚，其答案是2。

用这种连续减法找到除法余数的流程图可由下图给出。

为了使欧几里得算法的全部流程图完整，我们把上面形成余数的流程图代入到原先流程图的中右部的盒子中去。这种把一个算法向另一个代入是一种普遍的编电脑程序的步骤。上述寻求余数的算法是一个子程序的例子，也就是说，它是由主程序当做它的运算的一部分（通常是预先知道的）调用的算法。



当然，把数n简单地用n个×号来表示，在涉及大数时效率非常低，

这就是我们通常用更紧凑的诸如标准的（十进位）系统的原因。然而，我们在这里并不特别关心运算或记号的效率。我们所关心的是运算在原则上是否可被算法实行的问题。如果我们用一种数的记号是算法的东西，则用另一种记号也是算法的。两种情况中只有在细节上和复杂性上有差别。

欧几里得算法只是在整个数学中可找到的大量的经典算法步骤之一。可能出人意料的是，尽管算法的这一特殊例子有悠久的历史渊源，但一般算法概念的准确表达直到20世纪才诞生。事实上，这一概念的各种不同描述都是在20世纪30年代给出的。称作图灵机的概念是最直接的、最有说服力的，也是历史上最重要的。相当仔细地考查这些“机器”将是很适当的。

关于图灵“机”有一件事必须记在心里，就是说它是一段“抽象数学”，而不是一个物理对象。这一概念是由英国数学家，非凡的破码专家兼电脑科学的开山鼻祖阿伦·图灵在1935—1936年提出的（Turing 1937）。其目的是为了解决称为判决问题的一个范围广阔的问题。它是由伟大的德国数学家大卫·希尔伯特部分地于1900年巴黎国际数学家大会上（“希尔伯特第十问题”），更完整地于1928年波伦亚国际会议上提出的。希尔伯特不多不少地要求解决数学问题的一般算法步骤，或者不如讲，是否在原则上存在这样步骤的问题。希尔伯特还有一个要把数学置于无懈可击的坚固基础上的宏伟规划，其中公理和步骤法则一旦定下就不能变。但是在图灵完成其伟大的工作之际，这个规划已经遭受到由奥地利才气焕发的逻辑学家库尔特·哥德尔在1931年证明的令人吃惊的定理的粉碎性打击。我们将在第4章谈论哥德尔定理及其意义。图灵关心的希尔伯特问题（判决问题）超越出任何依据公理系统的特殊的数学形式。问题在于，是否存在能在原则上一个接一个地解决所有（属于某种适当定义的族的）数学问题的某种一般的机械步骤？

回答这一问题的部分困难在于决定什么叫作“机械过程”。这概念处于当时正常的数学概念之外。为了掌握它，图灵设想如何才能把“机器”的概念表达出来，它的动作被分解成基本的步骤。这一些似乎是清楚的，图灵也把人脑当成在他意义上的“机器”的例子。这样，由人类数学家在解决数学问题时进行的任何活动，都可以被冠以“机械步骤”之名。

虽然这一有关人类思维的观点似乎对于图灵发展他的极重要概念很有价值，我们却绝没有必要去附和它。的确，图灵在把机械过程的含义弄精确时，向我们展示出存在一些完好定义的数学运算，在任何通常的意义上，都不能被称为机械的！图灵本人的这一方面工作现在可间接地为我们提供了他自己有关精神现象性质观点的漏洞。这个事实也许不无讽刺意味。然而，这不是我们此刻所关心的。我们首先要弄清图灵心目中的机械过程究竟是什么。

图灵概念

我们设想实现某种（可以有限地定义的）计算步骤的一台仪器。这样仪器会采取什么样的一般形式呢？我们必须准备得理想化一些，而且不必为实用性过分担心：我们是真正地考虑一台数学上理想化的“机器”。我们要求该仪器具有有限（虽然也许非常大的）数目的不同可能态的分立集合。我们把这些称作仪器的内态。但是我们不限制该仪器在原则上要实现的计算的尺度。回顾一下上述的欧几里得算法。在原则上不存在被该算法作用的数的大小的限制。不管这些数有多大，算法或者一般计算步骤都是同样的。对于非常大的数，该步骤的确要用非常长的时间，而且需要在数量可观的“粗纸”上面进行实际的计算。但是不管这些数有多大，该算法是指令的同一有限集合。

这样，虽然我们仪器只有有限个内态，它却能够处理大小不受限制的输入。此外，为了计算应允许该仪器调用无限的外存空间（我们的“粗纸”）；而且能够产生大小不受限制的输出。由于该仪器只有有限数目的不同的内态，不能指望它把所有外部数据和所有自己计算的结果“内化”。相反地，它必须只考察那些立即处理的数据部分或者早先的计算，然后对它们进行任何需要的运算。也许可在外存空间把那个运算的相关结果记下来，然后以一种精确决定的方式进行下一个阶段的运算。正是输入、计算空间和输出的无限性质告诉我们，我们正在考虑的仅仅是一种数学的理想化，而不是在实际上可真正建造的某种东西（图2.1）。但这是极其关键的理想化。对于大多数实用目的而言，当代电脑技术的奇迹为我们提供了无限的电子存储仪器。

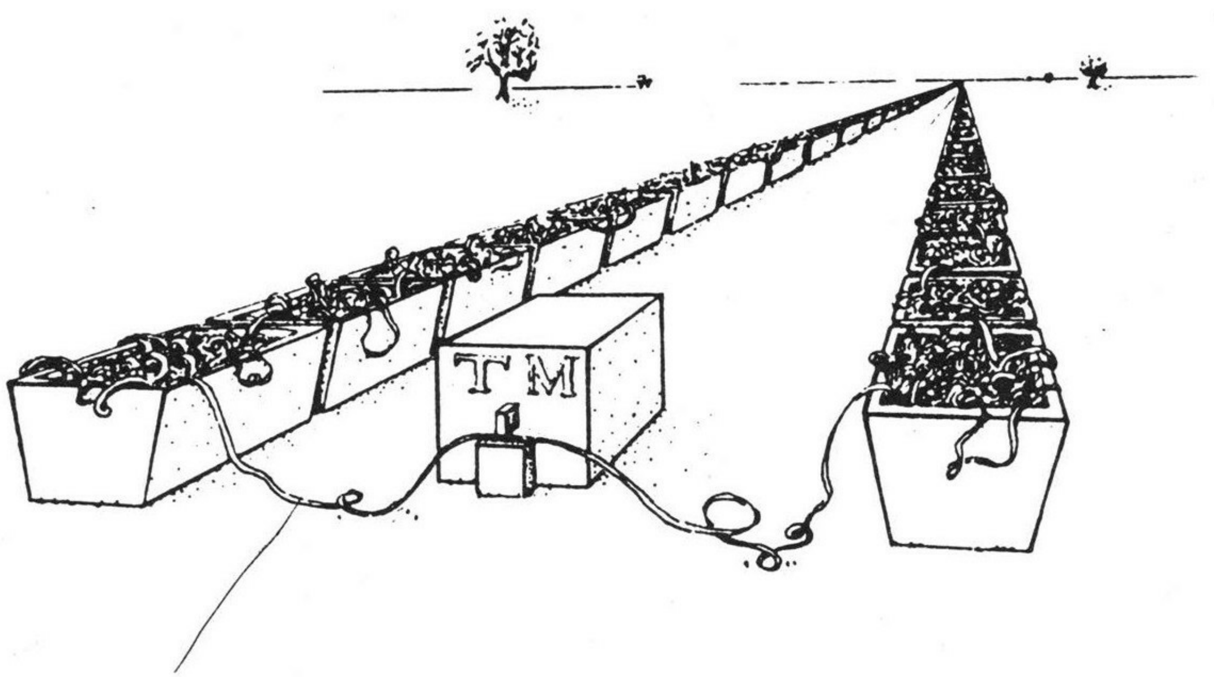


图2.1　一台严格的图灵机需要无限的磁带

事实上，在上述讨论中称为“外部的”存储空间的类型，可实际上被认为是现代电脑的内部工作的一部分。存储空间的某一确定部分是否被当作内部的或外部的，也许只是术语的问题。硬件和软件是划分“仪器”和“外部”的一种方法，内部可当成硬件，而外部为软件。我将不必拘泥于此，但是不管从哪个角度看，当代电子电脑是图灵的理想化的极好近似。

图灵是按照在上面做记号的“磁带”来描述其外部数据和存储空间的。一旦需要，仪器就会把该磁带调来“阅读”，而且作为其运算的一部分，磁带可前后移动。仪器可把记号放到需要的地方，可抹去旧的记号，允许同一磁带像外存（也就是“粗纸”）以及输入那样动作。因为在许多运算中，一个计算的中间结果起的作用正如同新的数据，所以事实上在“外存”和“输入”之间不做任何清楚的区分也许是有益的。我们记得在欧几里得算法中，不断地用计算不同阶段的结果去取代原先的输入（数A和B）。类似地，这同一磁带可被用作最后输出（也就是“答案”）。只必须进行进一步的计算，该磁带就会穿过该仪器而不断地前后移动。当计算最后完成时，仪器就停止，而计算的答案会在停于仪器一边的磁带的部分上显示出来。为了确定起见，我们假定，答案总是在左边显示，而输入的所有数据以及要解的问题的详细说明总是由右边进去。

让我们有限的仪器前后移动一条潜在的无穷长的磁带，我自己觉得有点不舒服。不管其材料是多么轻，移动无限长的磁带是非常困难的！相反地，我宁愿把磁带设想成代表某一外部环境，我们有限的仪器可以通过这环境进行移动。（当然用现代电子学，既不需要“磁带”也不需要“仪器”做实际的、通常物理意义上的“运动”，但是这种“运动”是一种描述事体的便利方法。）依此观点，该仪器完全是从这个环境接受它的输入。它把环境当成它的“粗纸”。最后将其输出在这同一个环境中写出。

在图灵的描述中，“磁带”是由方格的线性序列所组成，该序列在两个方向上都是无限的。在磁带的每一方格中或者是空白的或者包含有一个单独的记号 [5] 。我们可利用有记号或者没有记号的方格来阐释，我们的环境（也就是磁带）可允许被细分并按照分立（和连续相反的）元素来描述。希望仪器以一种可靠并绝对确定的方式工作，这似乎是合情理的要做的事。然而，我们允许该“环境”是（潜在地）无限的，把这作为我们使用的数学理想化的特征。但是，在任何特殊的情形下，输入、计算和输出必须总是有限的。这样，虽然可以取无限长的磁带，但是在它上面只应该有有限数目的实在的记号。磁带在每一个方向的一定点以外必须是空白的。

我们用符号“0”来表示空白方格，用符号“1”来代表记号方格，例如：

……	0	0	0	1	1	1	1	0	1	0	0	1	1	1	0	0	1	0	0	1	0	1	1	0	1	0	0	……

我们要求该仪器“读”此磁带，并假定它在一个时刻读一个方格，在每一步运算后向右或向左移动一个方格。这里没有损失任何涉及的一般性。可以容易地由另一台一次只读和移动一个方格的仪器去仿照一台一次可读n个方格或者一次可移动k个方格的仪器。k方格的移动可由一个方格的k次移动来积累，而存储一个方格上的n种记号的行为正和一次读n个方格一样。

这样的一台仪器在细节上可做什么呢？什么是我们描述成“机械的”东西作用的最一般的方式呢？我们记得该仪器的内态在数目上是有

限的。除了这种有限性之外，我们所需要知道的一切是该仪器的行为完全被其内态和输入所确定。我们已把输入简化成只是两个符号“0”或“1”之中的一个。仪器的初态和这一输入一给定，它就完全确定地运行；它把自己的内态改变成某种其他（或可能是同样的）内态，它用同样的或不同的符号0或1来取代它刚读过的0或1；它向右或向左移动一个方格；最后它决定是继续还是终止计算并停机。

为了以明白的方式定义该仪器的运算。我们首先，譬如讲用标号0，1，2，3，4，5，……来为不同的内态编号；那么，用一张显明的代换表可以完全指定该仪器或图灵机的运行，譬如：

$$0\mathbf{0} \rightarrow 0\mathbf{0}_R$$

$$0\mathbf{1} \rightarrow 13\mathbf{1}_L$$

$$1\mathbf{0} \rightarrow 65\mathbf{1}_R$$

$$1\mathbf{1} \rightarrow 1\mathbf{0}_R$$

$$2\mathbf{0} \rightarrow 0\mathbf{1}\text{R STOP}$$

$$2\mathbf{1} \rightarrow 66\mathbf{1}_L$$

$$3\mathbf{0} \rightarrow 37\mathbf{0}_R$$

$$\vdots \quad \vdots$$

$$210\mathbf{0} \rightarrow 3\mathbf{1}_L$$

$$\vdots \quad \vdots$$

258**0**→0**0**R. STOP

259**0**→97**1** R

259**1**→0**0**R. STOP

箭头左边的大写的数字是仪器在阅读过程中磁带上的符号，仪器用右边中间的大写的数字来取代之。R告诉我们仪器要向右移动一个方格，而L告诉我们它要向左移动一个方格。（如果，正如图灵原先描述的那样，我们认为磁带而不是仪器在移动，那么我们必须将R解释成把磁带向左移动一个方格，而L为向右移动一个方格。）词STOP表示计算已经完成而且机器就要停止。特别是，第二条指令$01 \rightarrow 131_L$ 告诉我们，如果仪器内态为0而在磁带上读到1，则它应改变到内态13，不改变磁带上的1，并沿着磁带向左移一格。最后一条指令2591→0 R.STOP告诉我们，如果仪器处于态259而且在磁带上读到，那么它应被改变为态0，在磁带上抹去1而产生0，沿着磁带向右移一格，然后终止计算。

如果我们只用由0到1构成的符号，而不用数字0，1，2，3，4，5……来为内态编号的话，则就和上述磁带上记号的表示更一致。如果我们有选择的话，可简单地用一串n个1来标号态n，但这是低效率的。相反地，我们使用现在人们很熟悉的二进制数系：

$$0 \to 0,$$

$$1 \to 1,$$

$$2 \to 10,$$

$$3 \to 11,$$

$4 \rightarrow 100$，

$5 \rightarrow 101$，

$6 \rightarrow 110$，

$7 \rightarrow 111$，

$8 \rightarrow 1000$，

$9 \rightarrow 1001$，

$10 \rightarrow 1010$，

$11 \rightarrow 1011$，

$12 \rightarrow 1100$，等等

正如标准的（十进位）记数法一样，这里最右边的数字代表“个位”，但是紧靠在它前面的位数代表“二”而不是“十”。再前面的位数代表“四”而不是“百”，更前面的是“八”而不是“千”等。随着我们向左移动，每一接续的位数的值为接续的二的幂：1，2，4（=2×2），8（=2×2×2），16（=2×2×2×2），32（=2×2×2×2×2）等。（为了将来的其他目的，我们有时发现用二和十以外的基来表示自然数是有用的：

例如基数为三，则十进位数64就可被写成2101，现在每一位数都为三的幂：64=（2×3^3 ）$+3^2$ +1；参阅第4章139页的脚注。）

对上面图灵机的内态使用这种二进制记数法，则原先的指令表便写成：

$$0\mathbf{0}\rightarrow0\mathbf{0}\mathrm{R}$$

$$0\mathbf{1}\rightarrow1101\mathbf{1}\mathrm{L}$$

$$1\mathbf{0}\rightarrow1000001\mathbf{1}\mathrm{R}$$

11→10R

100→01STOP

101→10000101L

110→1001010R

⋮ ⋮

110100100→111L

⋮ ⋮

1000000101→00STOP

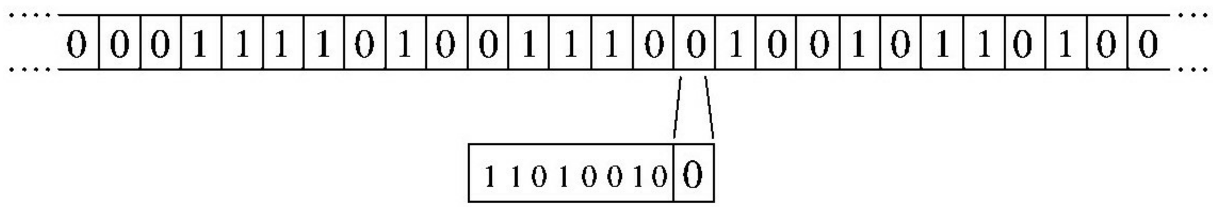
1000000110→1100001R

1000000111→00STOP

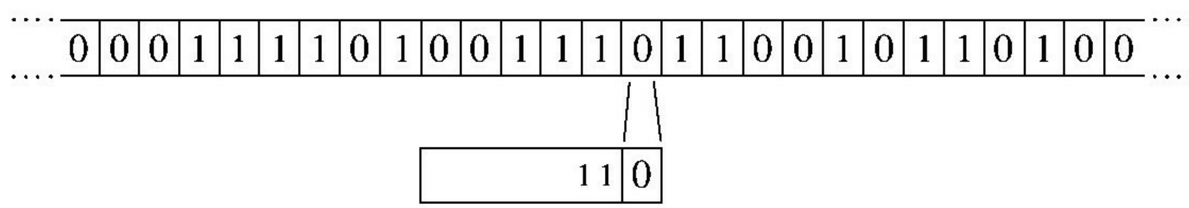
我还在上面把R.STOP简写成STOP，这是由于可以假定L.STOP从来不会发生，以使得计算的最后一步结果，作为答案的部分，总是显示在

仪器的左边。

现在假定我们的仪器处于由二进制序列1010010代表的特殊内态中，它处于计算的过程中，第44页给出了它的磁带，而且我们利用指令110100100→111L：



在磁带上被读的特殊位数（这里是位数“0”）由一个更大写的数字指示，符号串的左边表示内态。在由上面（多多少少是我随机造出的）部分的指定的图灵机例子中，读到的“0”会被“1”所取代，而内态变成“11”，然后仪器向左移动一格：



该仪器现在即将读另一个数字，它又是“0”。根据该表，它现在不改变这个“0”，但是其内态由“100101”所取代，而且沿着磁带向右移回一格。现在它读到“1”，而在表的下面某处又有如何进一步取代内态的指令，告诉它是否改变所读到的数，并向那个方向沿着磁带移动。它就用这种方式不断继续下去，直到达到STOP为止，在该处（在它向右再移一格之后）我们可以想象听到一声铃响，警告机器操作员计算完毕。

我们将假定机器总是从内态“0”开始，而且在阅读机左边的磁带原先是空白的。所有指令和数据都是在右边输进去。正如早先提到的，被提供的这些信息总是采用0和1的有限串的形式，后面跟的是空白带（也就是0）。当机器达到STOP时，计算的结果就出现在阅读机左边的磁带上。

由于我们希望能把数字数据当作输入的一部分，这样就需要有一种描述作为输入部分的通常的数（我这里是说自然数0，1，2，3，4，……）的方法。一种方法可以是简单地利用一串n个1代表数n（尽管这会给我们带来和自然数0相关的困难）：

1→1，2→11，3→111，4→1111，5→11111，等等。

这一初等的数系（相当非逻辑地）被称作一进制数系。那么符号“0”可用作不同的数之间的分隔手段。这种把数分隔开的手段是重要的，这是由于许多算法要作用到数的集合，而不仅仅是一个数上面。例如，对于欧几里得算法，我们的仪器要作用到一对数A和B上面。图灵机可以很容易地写下执行该算法的程序。作为一个练习，某些勤奋的读者也许介意去验证下面的一台图灵机（我将称它为EUC）的显明的描述，当应用到一对由0分隔的一进制数时，的确会执行欧几里得算法：

00→00R，01→11L，10→101R，11→11L，100→1010R，

101→110R，110→1000R，111→111R，1000→1000R，

1001→1010R，1010→1110L，1011→1101L，1100→1100L，

1101→11L，1110→1110L，1111→10001L，10000→10010L，

10001→10001L，10010→100R，10011→11L，

10100→00STOP，10101→10101R。

然而，任何读者在进行此事之前，从某种简单得多的东西，譬如图灵机UN+1开始将更为明智：

00→00R，01→11R，10→01STOP，11→01R。

它简单地把1加到一个一进制数上。为了检查UN+1刚好做到这点，

让我们想象，譬如讲把它应用到代表数4的磁带上去：

. . . 0000011110000 . . . 。

我们使仪器在开始时从某处向左为一些1。它处于内态0并且读到0。根据第一条指令，它仍保留为，向右移动一格，而且停在内态0上，在它遇到第一个1之前，它不断地这么进行并向右移动。然后第二条指令开始作用：它把1留下来不变并且再向右移动，但是现在处于内态1上。按照第四条指令，它停在内态1上，不改变这些1，一直向右移动，一直达到跟在这些1后面的第一个0为止。第三条指令接着告诉它把那个0改变成1，向右再移一步（记住STOP是表示R, STOP），然后停机。这样，另一个1已经加到这一串1上。正如所要求的，我们例子中的4已经变成了5。

作为更费神的练习，人们可以验证，下面所定义的机器UN×2，正如它所希望的，把一个一进制数加倍：

0**0**→0**0**R，0**1**→1**0**R，1**0**→10**1**L，1**1**→1**1**R，10**0**→11**0**R，

10**1**→100**0**R，11**0**→0**1**STOP，11**1**→11**1**R，100**0**→101**1**L，

100**1**→100**1**R，101**0**→10**1**L，101**1**→101**1**L。

在EUC的情形中，为了得到有关的概念，人们可用一些明显的数对譬如6和8来试验。正如以前一样，阅读机处于态0，并且初始时处在左边，而现在磁带从一开始的记号是这样的：

……0000000000011111101111111100000……

在许多步之后，图灵机停止，我们得到了具有如下记号的磁带：

……000011000000000000000……

而阅读机处于这些非零位数的右边。这样，所需的最大公约数正是所需要的（正确的）2。

要完全解释为何EUC（或者UN×2）在实际上完成所预想的，牵涉许多微妙之处，而且解释本身比机器更复杂，这是电脑程序的通常特征！（为了完全理解一个算法步骤为何能做到所预想的，牵涉到洞察。“洞察”本身是算法的吗？这是一个对我们以后颇为重要的问题。）我不想在这里为EUC或UN×2提供解释。真正做过检验的读者会发现，为了在所需的方案中把事情表达得更精密一些，我自作主张地对欧几里得算法作了一些不重要的变通。EUC的描述仍然有些复杂，对于11种不同的内态包含有22条基本指令，大部分复杂性是纯粹组织形式的。例如，可以看到在22条指令中，只有3条真正涉及在磁带上改变记号！（甚至对于UN×2用了12条指令，其中只有一半涉及改变记号。）

数据的二进制码

用一进制表示大数极端无效率。正如早先描述的，我们将相应地用二进制数系。然而，不能直接地把磁带当作二进制数来读。如果这样做的话，就没有办法告知一个数的二进制表示何时结束，以及无限个的序列是否代表右端开始的空白。我们需要某种终结一个数的二进制描述的记号。此外，我们还经常要输进几个数。正如欧几里得算法需要一对数[2] 那样。问题在于，我们不能把数之间的间隔和作为单独的一个数的二进制表示中的一部分的0或一串0区分开来。此外，我们或许在输入磁带中包括所有种类复杂的指令和数。为了克服这些困难，让我们采用一种我称之为收缩的步骤。按照该步骤，任何一串0或一串1（共有有限个）不是简单地被当作二进制数来读，而是用一串0，1，2，3等来取代。其做法是，第二个序列的每一数字就是在第一个序列中的连续的0之间的的1个数。例如序列

01000101101010110100011101010111100110

就可被取代成下面的形式：

010 0 01011010101011010 0 011101010111110 0110
| | | | | | | | | | | | | | | | | | |
1 0 0 1 2 1 1 2 1 0 0 3 1 1 4 0 2

我们现在可以把数2，3，4，……当作某种记号或指令来读。让我们把2简单地当作表示两个数之间间隔的“逗号”，而根据我们的愿望，3，4，5，……可以代表我们关心的各种指令或记号，诸如“负号”、“加”、“乘”“到具有下面号码的位置”，“将前述运算迭代如下多次”，等等。我们现在有了由高阶数分开的各种和的串。后者代表写成二进制的通常的数。这样上面可读成（“逗号”为2）：

（二进制数1001）逗号（二进制数11）逗号……

使用标准的阿拉伯数字“9”，“3”，“4”，“0”来写相应的二进制数1001，11，100，0，我们就得到整个序列：

9，3，4（指令3）3（指令4）0，

特别是，这一步骤给了我们一种简单地利用在结尾处逗号终结描述一个数的手段（并因此把它和在右边的无限长的空白带区分开来）。此外，它还使我们能对以二进制记号写成和的单独序列的自然数的任何有限序列编码。让我们看看在一特定情形下这是怎么进行的。例如，考虑序列

5，13，0，1，1，4，

在二进制记数法中这是

101，1101，0，1，1，100，

它可用扩展（也就是和上面收缩相反）的步骤在磁带上编码成

＜p＞

…00001001011010100101100110101011010101101000011000… 为了直截了当地得到这个码，我们可对原先的二进制数列作如下代换：

$$0 \to 0$$

$$1 \to 10$$

$$, \to 110$$

然后在两端加上无限个0。如果我们把它列出，就能更清楚地看出，如何把这个应用到上面的磁带上：

000010010110101001011001101011010110100011000

我将把这种数（的集合）的记号称为扩展二进制记号（这样，例如13的扩展二进制形式为1010010）。

关于这种编码还有最后一点必须提及。这只不过是个技巧，但是为了完备起见是需要的[3]。在自然数的二进制（或十进制）表示中，处于表式最左端的0是不“算”的，它通常可被略去，这里有些多余。例如00110010和110010是两个相同的二进制数（而0050和50为相同的十进制数）。这一多余可适合于数0本身，它也可写成000或00。一个空白的空间的确也应该逻辑地表示0！在通常的记号下这会导致巨大的混淆，但是它和上面刚描述的记号可相安无事。这样，在两个逗号之间的0可只写成两个连在一起的逗号（，），它在磁带上被编码成两对由单独的0隔开的11：

……001101100……

这样，上面的6个数的集合也可用二进制记号写成：

101，1101，1，1，100，

而且在磁带上可以扩展的二进制方式编码成：

……00001001011010100101101101011010110100011000……（有一个0已从我们以前的序列中略去）。

现在我们可以考虑让一台图灵机，譬如讲欧几里得算法，把它应用到以扩展二进制记数法写出的一对数上。例如，这一对数是我们早先考虑的6，8，不用以前用的：

……0000000000011111101111111100000……

而考虑6和8的二进制表示，也就是分别为110和1000。这一对为6，8，用二进制记数法也就是110，1000扩展后在磁带上编码成：

……00000101001101000011000000……

对于这一对特殊的数，并没有比一进制形式更紧凑。然而，譬如说我们取（十进制数）1583169和8610。在二进制记数法中它们是：

110000010100001000001，10000110100010，

这样，我们在磁带上把这一对编码成：

……0010100000010010000001000000101101000001010
010000100110……

只要用一行就可将其全部写出，而如果用一进制记数法的话，表示“1583169，8610”的磁带用这一整本书都写不下。

当数用扩展二进制记数法表示时，一台执行欧几里得算法的图灵机，如果需要的话，可以简单地把适当的一对在一进制和扩展二进制之间互相翻译的子程序算法接到EUC上去而得到。然而，由于一进制编数系统的低效率仍在“内部”存在，并且在仪器的迟缓以及需要大量的外部“粗纸”（它是磁带的右手部分）方面表现出来，实际上这是极其低效率的。可以给出全部用扩展二进制运算的、更有效率的、欧几里得算法的图灵机，但是它在这里对我们并无特别启发之处。

相反地，为了阐明如何使一台图灵机能对扩展二进制数运算，让我们尝试某种比欧几里得算法简单得多的东西，即是对一个自然数加1的

过程。这可由（我称之为XN+1的）图灵机来执行：

00→00R，01→11R，10→00R，11→101R，

100→110L，101→101R，110→01STOP，111→1000L，

1000→1011L，1001→1001L，

1010→1100R，1011→101R，

1101→1111R，1110→111R，1111→1110R，

某些勤奋的读者可把它应用到，譬如讲数167上去，以再次验证这一台图灵机在实际上做到了所预想的。这一个数的二进制表示可由下面的磁带给出：

……0000100100010101011000……

为了把1加到这个二进制数上，我们简单地找到最后的那个0，并把它改成1，然而用0来取代所有跟在后面的1。例如167+1=168在二进制记数法下写成：

10100111+1=10101000

这样，我们的“加1”图灵机应把前面的磁带用

……0000100100100001100000……

来取代，它的确做到了这一点。

我们注意到，甚至这种简单加1的非常基本的运算在用这种记数法时都会显得有些复杂，它使用了15条指令和8种不同的内态！由于在一进位系统中“加1”只是把1的串再延长1个而已，事情当然是简单得多。所以我们的机器UN+1更为基本，这一点也不奇怪。然而，对于非常大的数，由于所需的磁带非同寻常地长，UN+1就会极慢。而用更紧凑的扩展二进制记号运算的更复杂的机器XN+1就会更好。

作为旁白，我愿意指出对于扩展二进制比一进制图灵机显得更简单的一个运算，这就是乘二。在这里由

$$0\mathbf{0}\to 0\mathbf{0}\mathrm{R},\ 0\mathbf{1}\to 1\mathbf{0}\mathrm{R},\ 1\mathbf{0}\to 0\mathbf{1}\mathrm{R},\ 1\mathbf{1}\to 10\mathbf{0}\mathrm{R},\ 10\mathbf{0}\to 11\mathbf{1}\mathrm{R},$$

$$11\mathbf{0}\to 0\mathbf{1}\mathrm{STOP},$$

给出的图灵机XN×2能在扩展的二进制上实现这个运算，而前面描述的相应于一进制的机器UN×2就要复杂得多！

我们从这里得到了，关于图灵机在非常基础水平上可做到的一些事情的概念。正如所预料的，当进行某些复杂的运算时，它们会变得极为复杂。这种仪器的终极能力是什么呢？让我们在下面讨论这一个问题。

丘奇—图灵论题

人们一旦对建造简单的图灵机稍有一些熟悉，下面这些事实就很容易使他们感到满意。特殊的图灵机的确能执行各种基本的算术运算，诸如把两个数加到一起，或把它们相乘，或求一个数到另一个数的方次。显明地给出这种机器不是太啰唆的事，但是我不想在这里着手这么做。也可以提供其结果为一对自然数的运算，譬如带有余数的除法，或者其结果为任意大数目的数的有限集合。此外，可以这样地建造图灵机，即预先没有必要指定它要做何种运算，其运算的指令由磁带提供。需要实行的特定运算也许在某一阶段依赖于该机器在某个更早阶段的需要进行的某个计算的结果。（“如果那个计算的结果比某数大，则做这个；否则就做那个。”）一旦人们理解到，他可制造实现算术或简单的逻辑运算的图灵机，则很容易想象如何使之执行具有算法性质的更复杂的任务。在他们捣弄了好一阵之后，很容易坚信，这类机器的确能执行不管什么样的机械运算！在数学上，可以很合情合理地把机械运算定义为可被这样的一台机器所执行的运算。数学家用名词“算法”以及形容词“可计算的”、“递归的”和“有效的”来表示能由这类理论机器——图灵机实行的机械运算。一个步骤只要是足够清楚并且机械的，则可合理地相信能找到一台执行它的图灵机。这正是我们（也就是图灵）引进图灵机概念本身的初步讨论的全部要点。

另一方面，人们仍会感到，这些机器的设计不必这么局限。初看起来，只允许仪器在一个时刻读一个二进制数字（0或1），并且一次沿着一个单独的一维磁带只移动一格似乎是限制。为什么不允许大数目或相互联结的阅读机一下子跑过4条或5条甚至1000条分开的磁带呢？为什么不允许0和1的方格的整个平面（或者甚至一个三维的阵列），而坚持只用一维的磁带呢？为什么不允许从某种更复杂的计数系统或字母来的其他符号呢？事实上，虽然这些改变中的一些会对运算的经济性造成一定程度的不同（正如允许用多于一条的磁带一定会是这种情形那样），但是所有这一切都不会对我们在原则上要得到的东西造成丝毫的影响。即使我们在所有这些方面一下子推广该定义，这种归于“算法”的名下（或“计算”、“有效步骤”或“递归运算”）所实现的运算种类刚好和推广以前的完全相同！

我们可以看到，没有必要有多余一条的磁带，只要该仪器需要时总能在给定的磁带不断地找到新的空白。为此，也许必须不断地把数据从磁带的一处往另一处调度。这也许是“低效率的”，但是它不限制在原则上可以得到的极限 [4] 。类似地，利用多于一台并行动作的图灵机——这在近年来由于尝试更精密地模仿人脑而变得很时髦——不能在原则上得到任何新东西（虽然在某种情形下可改善动作的速度）。拥有两台分开的、不直接相互联络的仪器并不比两台相互联络的得到更多，而且如果它们联络，则实际上只不过是一台单独的仪器！

关于图灵的对于一维磁带的限制能说些什么呢？如果我们认为该磁带代表“环境”，也许宁愿把它当作一个平面或许一个三维空间，而不当做一维磁带。一个平面似乎比一维磁带更接近于一个“流程图”（正如在上面对欧几里得算法运算的描述）所需要的 [6] 。然而，在原则上以“一维的”形式（也就是利用流程图的通常术语来描述）写出流程图的运行并没有困难。在二维平面上显示只是为了我们的便利和容易理解，它对原则上能得到的并没造成什么影响。人们总能把一个二维平面上甚至三维空间上的一个记号或对象的地址，直截了当地在一维磁带上编码。（事实上，使用一个二维平面完全等效于用两条磁带。这两条磁带提供为在二维平面上指明一点所需的两个“坐标”；正如3条磁带可作为一个三维空间的一点的“坐标”一样。）这一维的编码又可能是“低效率的”，但是它在原则上不限制我们的目标。

尽管所有这一切，我们仍然可心存质疑，图灵机的概念是否真的和

我们希望叫作“机械的”每一逻辑或数学运算相统一。在图灵写他的开创性文章的时候，这一点比今天模糊得多，所以他觉得有必要把情形解释得更清楚一些。图灵的严密论证从以下事实得到了额外的支持，这就是美国逻辑学家阿伦佐·丘奇（在S.C.克林的协助下）完全独立地（并实际上稍早一些）提出了一种方案，也是旨在解决希尔伯特的判定问题的λ演算。尽管它不如图灵的那么明显地为一种完全广泛的机械的方案，但在数学结构上的极端经济性方面有些优点。（我将在本章的结尾描述丘奇的杰出的计算。）还存在其他一些解决希尔伯特问题的和图灵相独立的设想（见*Gandy 1988*），尤其是波兰——美国逻辑学家爱弥尔·波斯特的设想（比图灵稍晚些，但其思想与图灵比和丘奇更相像许多）。所有这些方案很快就被证明是完全等效的。这就给现在称作丘奇—图灵论题的观点增加了许多分量，即图灵机（或等效的）概念实际在数学上定义了我们认为是算法（或有效或递归或机械的）步骤的东西。现在，高速电脑已变成我们生活中如此熟悉的部分，很少人似乎觉得有必要去问这些论题的原始形式。相反地，已有不少人转去注意真正的物理系统（假定包括人脑）——精确服从物理定律的东西——是否能够执行比图灵机更多、更少或刚好一样多的逻辑和数学运算。我本人是非常喜欢丘奇—图灵论题的原先的数学形式。另一方面，它和实在物理系统的行为的关系是我们以后在本书主要关注的另外一个单独的问题。

不同于自然数的数

我们在上述的讨论中考虑了自然数的运算，并且注意到了这一显著的事实，即尽管每台图灵机只有固定的有限数目的不同内态，它却可能处理任意大小的自然数。然而，人们经常需要使用比这更复杂的其他种类的数，例如负数、分数或无理数。图灵机可以容易地处理负数和分数（例如像-597/26的数），而且我们可取任意大的分母和分子。我们所要做的全部是对“-”和“/”作适当的编码。这可容易地利用早先描述的扩展二进制记数法做到（例如，“3”表示“-”以及“4”表示“/”，它们分别在扩展二进制记数法中编码成1110和11110）。人们就是这样地按照自然数的有限集合来处理负数和分数的。这样，就可计算性的一般问题而言，它们没有告诉我们什么新的东西。

类似地，由于长度不受限制的有尽小数表式仅仅是分数的特殊情形，它们并没给我们带来什么新问题。例如，无理数π的由3.14159265

给出的有尽小数近似，简单地就是分数314159265/100000000。然而，无尽小数表式，譬如完全无尽展开

π=3.14159265358979……

出现了一定的困难。严格地讲，无论是图灵机的输入或者输出都不是无尽小数。人们也许会想到，我们可以找到一台图灵机，在其输出磁带上产生由π的小数展开的、所有的一个接一个位数3，1，4，1，5，9，……，我们就简单地让机器一直开下去好了。但这对于一台图灵机来讲，是不允许的。我们必须等待机器停了以后（铃声响过！）才允许去检查输出。只要机器还没有到达停止命令，其输出就可能要遭受到改变，所以不能对它信任。另一方面，在它到达停止时，其输出必须是有限的。

然而，存在一种合法地使图灵机以与此非常类似的方法，一个跟着一个地产生数字的步骤。如果我们希望展开一个无尽小数，譬如讲π，我们可以让一台图灵机作用于0上以产生整数部分3；然后使机器作用到1上，产生第一小数位1；然后使其作用于2上，产生第二小数位4；然后作用于3上，产生1，这样不断地下去。事实上，一定存在在这个意义上产生π的全部小数展开的图灵机，尽管要把它明显地造出来颇费一点周折。类似的评论也适用于许多其他无理数，譬如$\sqrt{2}$ =1.414213562……。然而，正如在下一章将要看到的，人们发现有些无理数（非常引人注目地）根本不能由任何图灵机产生。能以这种方式产生的数叫作可计算的（*Turing 1937* ）。那些不能的（实际上是绝大多数！）是叫作不可计算的。我们将在后面的章节中回到这件事体及有关的问题上来。按照物理理论，用可计算的数学结构能否足够地描述实在的物理对象（也就是人脑），是我们要关心的问题。

一般地讲，可计算性是数学中的一个重要问题。人们不应该将其当成只适合于这类数的事体。人们可有直接作用于诸如代数或三角的数学公式上的图灵机，或可进行微积分的形式运算的图灵机。人们所需的一切是某种准确地把所有涉及的数学符号编码成0和1序列的形式，然后再利用图灵机的概念。这毕竟是图灵在着手解决判定问题时心里所想的，即寻求回答具有一般性质的数学问题的算法步骤。我们将很快地回到这上面来。

通用图灵机

我还未描述通用图灵机的概念。虽然其细节是复杂的，但是它背后的原则并不十分复杂。它的基本思想是把任意一台图灵机T的指令的表编码成在磁带上表示成0和1的串。然后这段磁带被当作某一台特殊的被称作通用图灵机U的输入的开始部分，接着这台机器正如T所要进行的那样，作用于输入的余下部分。通用图灵机是万有的模仿者。“磁带”的开始部分赋予该通用机器U需要用以准确模拟任何给定机器T的全部信息！

为了了解这是如何进行的，我们首先需要一种给图灵机编号的系统方式。考虑定义某个特殊的，譬如讲在前面描述的图灵机的一个指令表。我们必须按照某种准确的方案把这表编码成0和1的串。我们可借助于以前采用的“收缩”步骤来办到。因为，如果我们用数2，3，4，5和6来分别代表符号R、L、STOP、箭头（→）以及逗点，那么我们就可以用110、1110、11110、111110以及1111110的收缩把它们编码。这样，出现在该表中的这些符号实际的串可以采用分别被编码成0和10的位数0和1。由于在该图灵机的表中，在二进制记数的结尾大写的数的位置足以把大写的0和1从其他小写的阿拉伯数字中区分开来，所以我们不需要用不同的记号。这样，110将被读成二进制数1101，而在磁带上被编码成1010010。特别是，0读作00，它可毫不含糊地被编码成0，或者作为被完全省略的符号。实际上我们可以不必对任何箭头或任何在它紧前头的符号进行编码，而依靠指令的数字顺序去标明哪些符号必须是什么。尽管在采用这个步骤时，在必要之处要提供一些额外的“虚”指令，以保证在这个顺序中没有缝隙。这样的做法具有相当好的经济性。（例如，图灵机XN+1没有告诉我们对110要做什么的命令，这是因为这条指令在机器运行时从不发生，所以我们应该插入一条“虚”指令，譬如讲1100→00R，它可合并到表中而不改变任何东西。类似地，我们应该把101→00R插入到XN×2中去。）若没有这些“虚的”，表中后面的指令的编码就会被糟蹋了。因为在结尾处的符号L或R足以把一条指令和另一条隔开，所以我们在每一指令中实际不需要逗号。因此，我们采用下面的编码：

0表示0或0，10表示1或1，110表示R，1110表示L，11110表示STOP。

作为一个例子，让我们为图灵机XN+1编码（插入指令1100→00R）。在去掉箭头和在它们紧前面的位数以及逗号之后。我们得到：

00R　11R　00R　101R　110L　101R　01STOP

1000L　1011L　1001L　1100R　101R　00R　1111R

111R　1110R

为了和早先说的相一致，我们可以去掉每一个00，并把每一个01简单地用1来取代，这样得到：

R11RR101R110L101R1STOP1000L1011L1

001L1100R101RR1111R111R1110R

如下是在磁带上的相应的码：

1101010110110100101101010011101001

0110101111010000111010010101110100

0101110101000110100101101101010101

01101010101101010100110

我们总是可以把开始的110（以及它之前的无限的空白磁带）删去。由于它表示00R，这代表开头的指令00→00R。而我已隐含地把它当作所有图灵机共有的。这样仪器可从磁带记号左边任意远的地方向右跑到第一个记号为止。而且，由于所有图灵机都应该把它们的描述用最后的110结束（因为它们所有都用R、L或STOP来结束），所以我们也

可把它（以及假想跟在后面的0的无限序列）删去。这可以算作两个小节约。所得到的二进制数是该图灵机的号码，它在XN+1的情况下为：

1010110110100101101010011101001011
0101111010000111010010101110100010
1110101000110100101101101010101011
01010101101010100。

这一特殊的数在标准十进制记数法中为：

450813704461563958982113775643437908.

我们有时不严格地把号码为n的图灵机称为第n台图灵机，并用T_n来表示。这样，XN+1是第450813704461563958982113775643437908台图灵机！

我们必须顺着这图灵机的“表”走这么远，才找到一台甚至只进行如此平凡的（在扩展二进制记数法中）对自然数加一的运算，这真使人印象深刻！（尽管在我的编码中还可以有很少的改善余地，但我认为自己进行得相当有效率。）实际存在某些更低号码的有趣的图灵机。例如，UN+1的二进制号码为：

101011010111101010

它只是十进制的177642！这样，只不过是把一个附加的1加到序列1的尾巴上的特别平凡的图灵机UN+1是第177642台图灵机。为了好奇的原因，我们可以注意在任一种进位制中“乘二”是在图灵机表中这两个号码之间的某处。我们找到XN×2的号码为10389728107，而UN×2的号码为

14929234

20919872026917547669.

人们从这些号码的大小，也许会毫不奇怪地发现，绝大多数的自然数根本不是可工作的图灵机的号码。现在我们根据这种编号把最先的13台图灵机列出来：

T_0: 00→00R，01→00R，

T_1: 00→00R，01→00L，

T_2: 00→00R，01→01R，

T_3: 00→00R，01→00STOP，

T_4: 00→00R，01→10R，

T_5: 00→00R，01→01L，

T_6: 00→00R，01→00R， 10→00R，

T_7: 00→00R，01→？？？，

T_8: 00→00R，01→100R，

T_9: 00→00R，01→10L，

T_{10}: 00→00R，01→11R，

T_{11}: 00→00R，01→01STOP，

T_{12}: 00→00R，01→00R， 10→00R。

其中，T_0 简单地就是向右移动并且抹去它所遇到的每一件东西，永不停止并永不往回退。机器T_1 最终得到同样的效应。但它是以更笨拙的方法，在它抹去磁带上的每个记号后再往后跳回。机器T_2 也和机器T_0 一样无限地向右移动，但是它更有礼貌，简单地让磁带上的每一件东西原封不动。由于它们中没有一台会停下，所以没有一台可以合格地被称为图灵机。T_3 是第一台可敬的机器，它的确是在改变第一个

（最左边）的1为0后便谦虚地停止。

T_4 遭遇了严重的问题。它在磁带上找到第一个1后就进入了一个没有列表的内态，所以它没有下一步要做什么的指令。T_8、T_9 和T_{10} 遇到同样的问题。T_7 的困难甚至更基本，把它编码的0和1的串涉及5个接续的1的序列：110111110。对于这种序列不存在任何解释，所以只要它在磁带上发现第一个1就被绊住。（我把T_7 或其他任何机器T_n ，它的n的二进制展开包含多于4个1的序列称为不是正确指明的。）机器T_5 、T_6 和T_{12} 遭遇到和T_0 、T_1 和T_2 类似的问题。它们简单地、无限地、永远不停地跑下去。所有T_0、T_1、T_2、T_4、T_5、T_6、T_7、T_8、T_9、T_{10} 和T_{12} 都是伪品！只有T_3 和T_{11} 是可工作的，但不是非常有趣的图灵机。T_{11} 甚至比T_3 更谦虚，它在第一次遇到1时就停止，并且没有改变任何东西！

我们应该注意到，在表中还有一个多余。由于T_6 和T_{12} 从未进入内态1，机器T_{12} 和T_6 等同，并在行为上和T_0 等同。无论这个多余还是表中的图灵机伪品我们都不必为之烦恼。人们的确可以改变编码以摆脱许多伪品和大大减少重复。所有这些都是以使我们可怜的通用图灵机变得更复杂作为代价。通用图灵机必须把所读到的号码n解码并假装成图灵机T_n 。如果我们可以把所有伪品（或者多余量）取走，这还是值得做的。但是，我们很快就会看到，这是不可能的！这样，我们就不触动我们的编码好了。

例如，可方便地把具有

……0001101110010000……

接续记号的磁带解释成某个数字的二进制表示。我们记得0在两端会无限地继续下去，但是只有有限个1。我还假定1的数目为非零（也就是说至少有一个1）。我们可以选择去读在第一个和最后一个1（包括在内）之中的有限的符号串，在上述的情况是为一自然数的二进制写法：

110111001，

它在十进制表示中为441。然而，这一过程只能给我们奇数（其二

进制表示以1结尾的数）。而我们要能表示所有的自然数。这样，我们采取移走最后的1的简单方案（这个1仅仅被当作表示这一程序的终止记号），而把余下来的当成二进制数来读 [5] 。因此，对于上述的例子，我们有二进制数：

11011100，

它是十进制的220。这个步骤具有零也用磁带上的记号代表的好处，也就是：

……0000001000000……

我们考虑图灵机T_n 对我们从右边提供给它的磁带上（有限的）0和1的串的作用。根据上面给出的方案，可方便地把这串也考虑作某一个数，譬如m的二进制代表。我们假定，机器T_n 在进行了一系列的步骤后最终到达停止（即到达STOP）。现在机器在左边产生的二进制数串是该计算的答案。让我们也以同样方式把这当作，譬如是p的二进制代表来读。我们把表达当第n台图灵机作用到m上时产生p的关系写成：

T_n （m）=p.

现在，以稍微不同的方式看这一关系。我们把它认为是一种应用于一对数n和m以得到数p的一个特别运算。（这样，若给定两个数n和m，视第n台图灵机对m作用的结果而得出p。）这一特别运算是一个完全算法的步骤。所以它可由一台特殊的图灵机U来执行。也就是说，U作用到一对（n, m）上产生p。由于机器U必须作用于n和m两者以产生单独结果p，我们需要某种把一对（n, m）编码到一条磁带上的方法。为此，我们可假定n以通常二进制记数法写出并紧接着以序列111110终结。（我们记得，任一台正确指明的图灵机的二进制数都是仅仅由0，10，110，1110和11110组成的序列，因此它不包含比4个1更多的序列。这样，如果T_n 是正确指明的机器，则111110的发生的确表明数n的描述已终结。）按照我们上面的规定，跟着它的每一件东西简单地是代表m的磁带（也就是，紧跟二进制数m的是1000……）。这样，这第二个部分简单地就是T_n 假设要作用的磁带。

作为一个例子，如果我们取n=11和m=6当做U要作用的磁带，其记

号序列为：

……00010111111011010000……

这是由以下

……0000（开始的空白带）

1011（11的二进制表示）

111110（终结n）

110（6的二进制表示）

10000……（余下的磁带）

组成的。

在T_n 作用到m上的运算的每一接续的步骤，图灵机U要做的是去考察n的表达式中的接续数位的结构，以使得在m的数位（也就是T_n 的磁带）上可进行适当的代换。在原则上（虽然在实践中肯定很繁琐）不难看到人们实际如何建造这样的一台机器。它本身的指令表会简单地提供一种，在每一阶段读到被编码到数n中的“表”中，应用到m给出的磁带的位数时，合适元素的手段。肯定在m和n的数位之间要有许多前前后后的进退，其过程会极为缓慢。尽管如此，一定能提供出这台机器的指令表，而我们把它称为通用图灵机。把该机器对一对数n和m的作用表为U（n, m），我们得到：

U（n, m）=T_n （m）.

这里T_n 是一台正确指明的图灵机 [6] 。当首先为U提供数n时，它准确地模拟第n台图灵机！

因为U为一台图灵机，它自身也必须有一号码；也就是说，我们有：

U=T_u ,

此处号码u待定。u究竟是多少呢？事实上，我们可以准确地给出：

u=7244855335339317577198395039615711237952360672556559631108144796606505059404241090310483613632359365644443458382226883278767626556144692814117715017842551707554085657689753346356942478488597046934725739988582283827795294683460521061169835945938791885546326440925525505820555989451890716537414896033096753020431553625034984529832320651583047664142130708819329717234151056980262734686429921838172157333482823073453713421475059740345184372359593090640024321077342178851492760750481673375621725734645227310544829807849651269887889645697609066342044779890219144379328300194935709639217039048332708825962013017737272027186259199144282754374223513556751340842222998893744105343054710443686958764051781280194375308138706399427728231564252892375145654438990527807932411448261423572861931183326106561227555318102075110853376338060310823616750456358521642148695423471874264375444287900624858270912404220765387542644541334517485662915742999095026230097337381377241621727477236102067868540028935660856968226201419824862169890260913094029857060017430067008689675903447341741278742558120154936639389969058177385916540553567040928213322216314109787108145997866959970450968184190629944365601514549048809220844800348224920773040304318842989939313526688234966210194716191070146196852319284748203449589770955356110702758174873332729667899879847328409819076485127263100174016678736347760585724503696443489799203448999745566240293748766883975140445166570775006051388399166881407254554466522205072426239237921152531816251253630509317286314220040645713052758023076651833519956891397481375049264296050100136519801869 45639498

（或者至少是这等数量级的其他的可能性）。这个数无疑是极其巨大，但是我似乎没有办法使它变小许多。虽然我的图灵机编码步骤和号码指定是相当合理和简单的，但是在一台实际的通用图灵机的编码中仍然不可避免地导向这么大的一个数[7]。

我曾说过，实际上所有现代通用电脑都是通用图灵机。我并不是说，这种电脑的逻辑设计必须在根本上和我刚刚给出的通用图灵机的描述非常相似。其要点可以简述为，首先为任一台通用图灵机提供一段适当的程序（输入磁带的开始部分）可使它模拟任何图灵机的行为！在上面的描述中，程序简单地采取单独的数（数n）的形式。但是，其他的步骤也是可能的，图灵原先方案就有许多种变化。事实上，在我自己的描述中，已经有些偏离图灵的原型。但是对于我们当前的目的，这些差别中没有一个是重要的。

希尔伯特问题的不可解性

我们现在回到当初图灵提出其观念的目的，即解决希尔伯特的范围广泛的判定问题：是否存在某种回答属于某一广泛的，但是定义得很好的集合的所有数学问题的机械步骤？图灵发现，他可以把这个问题重述成他的形式，即决定把第n台图灵机作用于数m时实际上是否会停止的问题。该问题被称作停机问题。很容易建造一个指令表使该机器对于任何数m不停。（例如，正如上面的n=1或2或任何别的在所有地方都没有STOP指令的情形）。也有许多指令表，不管给予什么数它总停（例如n=11）；有些机器对某些数停，但对其他的数不停。人们可以公正地讲，如果一个想象中的算法永远不停地算下去，则并没有什么用处。那根本不够格被称作算法。所以一个重要的问题是，决定T_n 应用在m时是否真正地给出答案！如果它不能（也就是该计算不停止），则就把它写成：

T

n

（m）=□。

[在这记号中还包括了如下情形，即图灵机在某一阶段由于它找不到合适的告诉其下一步要做什么的指令而遇到麻烦，正如上面考虑的伪品机器T_4 和T_7 。还有不幸的是，我们粗看起来似乎成功的机器T_3 现在也必须被归于伪品：T_3 （m）=□，这是因为T_3 作用的结果总是空白带，而为使计算的结果可赋予一个数，在输出上至少有一个1！然而，

由于机器T_{11} 产生了单独的1，所以它是合法的。这一输出是编号为0的磁带，所以对于一切m，我们都有T_{11}（m）=0。]

能够决定图灵机何时停止是数学中的一个重要问题。例如，考虑方程：

$$(x+1)^{w+3}+(y+1)^{w+3}=(z+1)^{w+3}。$$

（如果专门的数学方程使你忧虑，不要退缩！这一方程只不过是作为一个例子，没有必要详细地理解它。）这一特殊的方程和数学中著名的或许是最著名的未解决的问题相关。该问题是：存在任何满足这方程的自然数集合w, x，y, z吗？这个著名的称作“费马大定理”的陈述被伟大的17世纪法国数学家皮埃尔·德·费马（1601—1665）写在丢番图的《代数》一书空白的地方。费马宣布这方程永远不能被满足 [7] [8] 。虽然费马以律师作为职业（并且是笛卡儿的同时代人）。他却是那个时代最优秀的数学家。他宣称得到了这一断言的“真正美妙的证明”，但那里的空白太小写不下。可惜迄今为止既没有人能够重新证明之，也没有人能找到任何和费马断言相反的例子[8]！

很清楚，在给定了4个数（w, x，y, z）后，决定该方程是否成立是计算的事体。这样，我们可以想象让一台电脑的算法一个接一个地跑过所有的四数组，直到方程被满足时才停下。（我们已经看到，存在于一条单独磁带上，把数的有限集合以一种可计算方式编码成为一个单独的数的方法。这样，我们只要跟随着这些单独的数的自然顺序就能“跑遍”所有的四数组。）如果我们能够建立这个算法不停的事实，则我们就有了费马断言的证明。

可以用类似的办法把许多未解决的数学问题按图灵机停机问题来重述。“哥德巴赫猜想”即是这样的一个例子，它断言比2大的任何偶数都是两个素数之和 [9] 。决定给定的自然数是否为素数是一个算法步骤，由于人们只需要检验它是否能被比它小的数整除，所以这只是有限计算的事体。我们可以设计跑遍所有偶数6，8，10，12，14，……的一台图灵机，尝试把它们分成奇数的对的所有不同的方法：

6=3+3，8=3+5，10=3+7=5+5，12=5+7，

14=3+11=7+7，……

对于这样的每一个偶数检验并确认其能分成都为素数的某一对数。（我们显然不需要去检验除了2+2之外的偶的被加数对，由于除了2之外所有素数都是奇的。）只有当我们的机器达到一个由它分成的所有的任何一对数都不是素数对的偶数为止才停止。我们在这种情形就得到了哥德巴赫猜想的反例，也就是说一个（比2大的）偶数不是两个素数之和。这样，如果我们能够决定这台图灵机是否会停，我们也就有了判定哥德巴赫猜想真理性的方法。

这里自然地产生了这样的问题：我们如何判定任何特殊的图灵机（在得到特定输入时）会停止否？对于许多图灵机回答这个问题并不难：但是偶尔地，正如我们上面得到的，这答案会涉及一个杰出的数学问题的解决。这样，存在某种完全自动地回答一般问题，即停机问题的算法步骤吗？图灵指出这根本不存在。

他的论证的要点如下所述，我们首先假定，相反的，存在这样的一种算法[10]。那么必须存在某台图灵机H，它能“判定”第n台图灵机作用于数m时，最终是否停止。我们假定，如果它不停的话，其输出磁带编号为0，如果停的话为1：

$$H(n;m)=\begin{cases}0 \text{ 如果 } T_n(m)=\square \\ 1 \text{ 如果 } T_n(m) \text{ 停止。}\end{cases}$$

在这里人们可采取对通用图灵机U用过的同样规则给对（n，m）编码。然而，这会引起如下技术问题，对于某些数n（例如n=7），T_n 不是正确指定的，而且在磁带上记号111101不足以把n从m分开。为了排除这一个问题，让我们假定n是用扩展二进制记数法而不仅仅是二进制记数法来编码，而m正和以前一样用通常的二进制记数法。那么记号

110实际上将足以把n和m区分开来。在H（n；m）中用分号，而在U（n, m）中用逗号就是为了表明这个改变。

现在让我们想象一个无穷阵列，它列出所有可能的图灵机作用于所有可能的不同输入的所有输出。阵列的n行展现当第n台图灵机应用于不同的输入0，1，2，3，4，……时的输出：

m→ n↓	0	1	2	3	4	5	6	7	8	…
0	□	□	□	□	□	□	□	□	□	…
1	0	0	0	0	0	0	0	0	0	…
2	1	1	1	1	1	1	1	1	1	…
3	0	2	0	2	0	2	0	2	0	…
4	1	1	1	1	1	1	1	1	1	…
5	0	□	0	□	0	□	0	□	0	…
6	0	□	1	□	2	□	3	□	4	…
7	0	1	2	3	4	5	6	7	8	…
8	□	1	□	□	1	□	□	□	1	…
•	•				•				•	…
•	•				•				•	
•	•				•				•	
197	2	3	5	7	11	13	17	19	23	…
•	•				•				•	
•	•				•				•	
•	•				•				•	

我在此表中稍微搞了点欺骗，并且没有把图灵机按它们的实际编号

列出。由于所有n比11小的机器除了□外没有得到任何东西，而对于n=11只得到0，所以那样做的话就会得到一张一开始就显得过于枯燥的表。为了使此表一开始就显得更有趣，我假定已得到某种更有效得多的编码。事实上我只是相当随机地捏造这张表的元素，仅仅是为了给出有关它的外表的大体印象。

我不要求用某一个算法实际计算过这一个阵列。（事实上，正如我们很快就要看到的，不存在这样的算法。）我们仅仅是假想，真正的表不知怎么搞的已经摆在我们面前。如果我们试图计算这一个阵列，正是□的发生引起了困难。因为既然那些计算简单地一直永远算下去，我们也许弄不清什么时候把□放在某一位置上！

然而，如果我们允许使用假想的H，由于H会告诉我们□实际上在什么地方发生，我们就可以提供一种产生该表的计算步骤。但是相反的，我们用0来取代每一次□的发生，就这样利用H把□完全除去。这可由把计算H（n；m）放在T_n 对m作用之前而做到；然后只有如果H（n；m）=1时（也就是说，只有如果计算T_n （m）实际上给出一个答案时），我们才允许T_n 作用到m上，而如果H（n；m）=0（也就是如果T_n （m）=□），则简单地写为0。我们可把新的步骤（也就是把H（n；m）的作用放在T_n （m）之前得到的）写成：

T_n （m）×H（n；m）.

（我在这里使用数学运算顺序的普通习惯：在右边的先进行。请注意，我们在符号运算上有：□×0=0。）

现在这张表变成：

m→ / n↓	0	1	2	3	4	5	6	7	8	…
0	0	0	0	0	0	0	0	0	0	…
1	0	0	0	0	0	0	0	0	0	…
2	1	1	1	1	1	1	1	1	1	…
3	0	2	0	2	0	2	0	2	0	…
4	1	1	1	1	1	1	1	1	1	…
5	0	0	0	0	0	0	0	0	0	…
6	0	0	1	0	2	0	3	0	4	…
7	0	1	2	3	4	5	6	7	8	…
8	0	1	0	0	1	0	0	0	1	…
•	•				•				•	
•	•				•				•	
•	•				•				•	

我们注意到，假定H存在，该表的行由可计算的序列组成。（我用一个可计算序列表明一个其接连的值可由一个算法产生出来的一个无限序列；也就是存在一台图灵机，当它依序地应用于自然数m=0，1，2，3，4，5，……上时，就得到了这个序列的接续元素。）现在，我们从该表中可以注意到两个事实。首先自然数的每一可计算序列必须在它的行中出现在某处（也许出现好几回）。这个性质对于原先的带有□的表已经是真的。我们只不过是简单地加上一些行去取代“伪品”的图灵机（也就是至少产生一个□的那些）。其次，我们已经假定，图灵机H实

际上存在，该表已被计算地（也就是用某个确定的算法）由步骤T_n（m）×H（n；m）产生。换言之，存在某一台图灵机Q，当它作用于一对数（n；m）时就会在表中产生合适的元素。为此我们可以在Q的磁带上以和在H中一样的方式对n和m编码。我们得到：

$$Q(n;m)=T_n(m)\times H(n;m).$$

现在我们应用乔治·康托尔的“对角线方法”的天才的和强有力的技巧的变种。（下一章将看到康托尔对角线方法的原型。）考虑现在用粗体字标明的对角线元素：

0	0	0	0	0	0	0	0	0
0	**0**	0	0	0	0	0	0	0
1	1	**1**	1	1	1	1	1	1
0	2	0	**2**	0	2	0	2	0
1	1	1	1	**1**	1	1	1	1
0	0	0	0	0	**0**	0	0	0
0	0	1	0	2	0	**3**	0	4
0	1	2	3	4	5	6	**7**	8
0	1	0	0	1	0	0	0	**1**
·				·				·
·				·				·
·				·				·

这些元素提供了某一序列0，0，1，2，1，0，3，7，1，……，现

在把它的每一元素都加上1就得到

1，1，2，3，2，1，4，8，2，……

假设我们的表是计算地产生的，那么这就清楚地是一个可计算的步骤，而且它为我们提供了某一个新的可计算的序列。事实上为1+Q（n；n），也就是：

$1+T_n$（n）×H（n；n）

（由于对角线是令m等于n而得到的）。但是，我们的表包括了每一可计算的序列，所以我们新的序列必须在表中的某一行。然而，这是不可能的！由于我们新序列和第一行在第一元素处不同，和第二行在第二元素处不同，和第三行在第三元素处不同，等等。这是一个明显的冲突。正是这个冲突，建立了我们所要证明的，即在事实上图灵机H不存在！不存在决定一台图灵机将来停止与否的通用算法。

另一种重述这个论证的方法是注意到，在假定H存在时，对于算法1+Q（n；n）（对角线步骤！）存在某一图灵机号码，譬如k，这样我们有：

$1+T_n$（n）×H（n；n）$=T_k$（n）.

但是，如果我们在这个关系中代入n=k就得到：

$1+T_k$（k）×H（k；k）$=T_k$（k）.

因为如果T_k（k）停止，我们就得到了不可能的关系式：

$1+T_k$（k）$=T_k$（k），

[由于H（k；k）=1]，而如果T_k（k）不停止[这样H（k；k）=0]，我们有同样不协调的结果：

1+0=□，

所以前面的假定导致一个矛盾。

一台特定的图灵机是否停止是一个定义完好的数学问题（反过来，我们已经看到，各种有意义的数学问题可被重述成图灵机的停机问题）。这样，依靠显示不存在决定图灵机停机问题的算法，图灵（正如丘奇用自己十分不同的手段）指出，不存在决定数学问题的一般算法。希尔伯特判定问题没有解答！

这不是说，在任何个别的情形下，我们不可能决定某些特殊数学问题的真理或非真理；或者决定某一台给定的图灵机是否会停止。我们可以利用一些技巧或者仅仅是常识，就能在一定情况下决定这种问题。（例如，如果一台图灵机的程序表中不包括STOP指令。或者只包含STOP指令，那么常识就足以告诉我们它会不会停止！）但是，不存在一个对所有的数学问题，也不存在对所有图灵机以及所有它们可能作用的数都有效的一个算法。

我们似乎现在已经建立了，至少存在某些非决定的数学问题。然而，我们从未做过这种事！我们还没有展示过，存在某种特别尴尬的数时，它在某种绝对的意义上，不可能决定该机器是否停止。的确，正如我们很快就要看到的，情况刚好相反。我们关于单独问题的不可解性一点也没提，仅仅是说关于问题的族的算法的不可解性。在任何单独的情形下，答案或者为“是”或者为“非”，所以肯定存在一个决定那个特定情形的算法，也就是在它面临该问题时，依情况而定，会简单地讲“是”或“非”。当然，困难在于我们可能不知道用这些算法中的哪一个。那就是决定一个单独陈述而不是决定一族陈述的数学真理的问题。重要的是要意识到，算法本身不能决定数学真理。一个算法的成立总是必须由外界的手段来建立起来。

如何超越算法

我们在以后论及哥德尔定理时再回到决定数学陈述的真理性问题（参阅第4章）。我希望在此刻指出，图灵的论证比我迄今所暗示的更具建设性而更少负面性。我们肯定还没有展示出一台特殊的图灵机，它在某种绝对的意义上不能决定其是否停止的问题。的确，如果我们仔细地考察该论证就会发现，我们的步骤本身实际上已经隐含地告诉我们，对这利用图灵步骤建造的似乎“极端荒谬”的机器的答案！

我们看看这是怎么来的。假定我们有某一个算法，它有时有效地告诉我们什么时候一台图灵机将不停止。正如在上面概述的，图灵步骤将明白地展现了一台图灵机计算，对这计算那个特殊算法不能决定其是否停止。然而，在这样做的同时，它实际上使我们在这种情况下看到了答案！我们展现的特殊的图灵机计算的确不会停止。

为了仔细考查这是怎么引起的，假定我们具有一个这样有时有效的算法。正如以前那样，我们用H来标志这个算法（图灵机），但是现在允许该算法有时不能告诉我们一台图灵机在实际上将不停止：

$$H(n;m)=\begin{cases}0\text{或}\square\text{，如果}T_n(m)=\square\text{；}\\ 1\text{，}\quad\text{如果}T_n(m)\text{停止。}\end{cases}$$

这样，当T_n（m）=□时H（n；m）=□是一种可能性。实际上存在许多这种算法H（n；m）。[例如，只要T_n（m）一停止，H（n；m）就能简单地产生1，尽管那个特殊的算法几乎没有什么实际的用处！]

除了不把所有的□用0来取代而留下一些以外，我们可以像上述那样仔细地顺着图灵的步骤。正如以前那样，对角线过程提供了对角线上的第n个元素

1+Tn（n）×H（n；n）.

[只要H（n；m）=□，我们就将得到一个□。注意□×□=□，1+□=□。]这是一个完好的计算，所以它是由某一台，譬如讲第n台图灵机得到的，而且现在我们确实有：

$1+T_n$（n）×H（n；m）=T_k（n）.

我们看第k个对角元素，也就是n=k，就会得到：

$1+T_k$（k）×H（k；k）=T_k（k）.

如果计算T_k（k）停止，我们就有了一个矛盾[由于假定只要T_k（k）停止H（k, k）就为1，则方程导致不协调性：1+T_k（k）=T_k（k）。所以T_k（k）不能停止，也就是：

T_k（k）=□。

但是，算法不能“知道”这个。因为如果该算法给出H（k, k）=0，则我们应该又导致矛盾（我们得到了在符号上不成立的关系：1+0=□）。

这样，如果我们能找到k，就将知道如何去建造击败我们知道其答案的算法的特别计算！我们怎么找到k呢？这是一项艰巨的工作。我们需要做的是仔细考察H（n；m）和T_n（m）的构造，然后仔细弄清1+T_n（n）×H（n；n）作为一台图灵机是如何动作的。我们发现这台图灵机的号码为k。要把这一切弄透彻肯定是复杂的。但它是可以办得到的[11]。由于这种复杂性，若不是因为我们为了击败H而特地制造T_k（k）的这个事实，我们对计算T_k（k）毫无兴趣！重要的是，我们有了定义很好的步骤，不管我们的H是哪一个，该步骤都能找到合适的k，使得T_k（k）击败H，因此这样我们可以比该算法做得更好。如果我们认为自己仅仅比算法更好些，也许会给我们带来一些小安慰！

事实上，该过程被定义得如此之好，以至于在给定的H下，我们可找到产生k的一个算法。这样，在我们过于得意之前必须意识到，由于这个算法事实上“知道”T_k（k）=□，所以它能改善[9] H，是不是？在上面提到一个算法时，用拟人化的术语“知道”是有助的。然而，该算法仅仅是跟随我们预先告诉它去跟随的法则，这难道不是我们在“知道”吗？或者我们自己，仅仅是在跟随我们的头脑的构造和我们的环境所预先编排我们去跟随的规则？这问题实在不只是简单的算法问题，而且是人们如何判断何为真何为伪的问题。这是我们必须重新讨论的中心问题。数学真理（及其非算法性质）的问题将在第4章考虑。现在我们至少对术语“算法”和“可计算性”的意义以及某些相关的问题已有些领略。

丘奇的λ计算法

可计算性的概念是一个非常重要和漂亮的数学观念。它又是相当近

代的，具有这样基本性质的事体进入数学的王国是20世纪30年代的事。这个观念已经渗透到数学的所有领域中去（虽然这一点确实是真的，但是大多数数学家通常不去忧虑可计算性的问题）。该观念的威力有一部分来自于这一个事实，即数学中一些定义得很好的运算实际上不是可计算的（例如图灵机的停机问题；第4章还可以看到其他例子）。因为如果不存在这种不可计算的事体，则可计算性的概念便没有多少数学的兴趣。数学家毕竟喜欢困惑的东西。让他们决定某些数学运算是否为可计算的可能是非常迷人的困惑。因为那个困惑的一般解答本身是不可计算的，这一点尤其迷人！

有一件事要弄清楚。可计算性是一个真正“绝对的”数学概念。它是一种抽象的观念，它完全超越按照我们描述的“图灵机”的任何特别实现之外。正如我在以前所评论的，我们不必对为表征图灵的天才而采用特别的手段的“磁带”和“内态”等赋予任何特别的意义。还有表达可计算性观念的其他方法，历史上最早的是美国逻辑学家阿伦佐·丘奇在斯蒂芬·C.克林协助下提出的杰出的“λ计算法”。丘奇的步骤和图灵的完全不同，并且更为抽象得多。事实上，在丘奇陈述他观念的形式中，在它们和任何可以称作“机械的”东西之间只有一点明显的连接。丘奇步骤背后的关键观念在其最本质上的确是抽象的，实际上丘奇把这步骤称为“抽象化”的一个数学运算。

不仅是因为丘奇方案强调可计算性是一个独立于计算机器的任何特别概念的数学观念，而且它阐明了在数学中抽象观念的威力，所以我感到值得花一点时间来简要地描述它。对数学观念不熟悉或者对这件事本身不感好奇的读者，在这一阶段可以跳到下一章去，这不会对论证的过程产生多少损失。尽管如此，这样的读者若愿意和我多忍受一阵会得到好处，并且能见证丘奇方案的某些魔术般的经济性（参见*Church 1941*）。

人们在此方案中关心的是，譬如由以下表示的对象的“宇宙”：

a, b，c, d，……，z, a'，b'，……，z'，a''，b''，……，

a'''，……，a''''，……，

其中每一元素代表一个数学运算或函数。（之所以用带撇的字母，只不过是便于无限地给出这些符号以表示这种函数。）这些函数的“自

变量”，即它们所作用的东西，是同一类型的其他东西。也就是函数。此外，一个这种函数作用于另一个函数的结果（或“值”）仍是一个函数。（在丘奇的系统中的确具有美妙的概念经济性。）这样，当我们写[12]

a=bc

时，我们是指函数b作用于函数c的结果为另一函数a。要在这个方案中表达两个或更多变量的函数的观念并没有困难。如果我们希望把f认为两个变量，譬如讲p和q的函数，我们可以简单地写：

（fp）q

（这是函数fp作用于q的结果）。对于三变量函数我们考虑：

（（fp）q）r，

等等。

让我们引进抽象化的有力的运算。为此我们使用希腊字母λ（拉姆达），而且直接再加上一个字母如x，以代表一个丘奇函数，我们把它当成“虚变量”。任何发生于紧跟在这后面的方括号内的表达式中的变量x是仅仅被当作一个“插口”，可以往里面代入任何跟在整个表达式后的任何东西。这样，如果我们写：

λx.[fx]，

我们是说，当它作用到譬如讲函数a上时，就产生结果fa。那就是：

（λx.[fx]）a=fa。

换言之，λx.[fx]简单地就是函数f，即：

λx.[fx]=f。

这里只用一点思维就够了。数学的一个美妙在于，初看起来是如此卖弄学识的、琐碎的东西，也是人们非常容易完全失去要点的东西。让

我们考虑从中学数学拿来的一个熟悉的例子。我们取函数f为对一个角度取正弦的三角运算，这样抽象的函数“sin”被定义为：

λx.[sinx]=sin。

（不必为何以“函数”x可当作一个角度而忧虑。我们很快就会看到数可被当成函数的某种方法；而一个角度只是一个数。）迄今为止的一切的确是相当无聊的。让我们设想，记号“sin”还没被发明，但是我们熟悉sinx的级数展开表达式：

$$x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \cdots,$$

然后我们可以定义：

$$\sin = \lambda x.\left[x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \cdots\right]。$$

请注意，我们甚至可以更简单地定义，譬如讲“六分之一立方”的运算，而它是没有标准的“函数”记号的：

$$Q = \lambda x.\left[\frac{1}{6}x^3\right],$$

而且我们发现，例如：

$$Q(a+1)=\frac{1}{6}(a+1)^3=\frac{1}{6}a^3+\frac{1}{2}a^2+\frac{1}{2}a+\frac{1}{6}\text{。}$$

从丘奇的基本函数运算简单构造的表达式对于现在的讨论更为贴切，例如：

λf.[f（f x）]。

这是一个函数，当它作用于另一函数，譬如讲g时，产生g两次迭代地作用于x上的函数，也就是：

（λf.[f（f x）]）g=g（gx）。

我们也可以首先“抽象化走”x以得到：

λf.[λx.[f（f x）]]，

此式可以缩写成：

λfx[f（fx）]。

这是当作用于g时产生“g被迭代两次”的函数。事实上，这正是丘奇将其和自然数2相等同的函数。

2 =λf x.[f（fx）]，

这样，（2g）y=g（gy）。它类似地定义：

3 =λf x.[f（f（fx））]，

4 =λfx[f（f（f（fx）））]，等等，

以及

1 =λf.[fx]，0=λfx.[x]。

丘奇的“2”真的更像“两次”，它的“3”是“3次”等。这样3在一个函数f上的作用也就是3f，是“把f迭代3次”的运算。因此，3f在y上的作用是（3f）y=f（f（f（y）））。

让我们看一个非常简单的算术运算，也就是如何把加到一个数上的运算在丘奇方案中表达出来。定义

S=λabc.[b（（ab）c）]。

为了阐明S的确简单地把加到用丘奇记号表示的一个数上，让我们做这样的验算：

$$S\mathbf{3}=\lambda abc.[b((ab)c)]\mathbf{3}=\lambda bc.[b((3b)c)]=\lambda bc.[b(b(b(bc)))]=\mathbf{4},$$

这是由于（3b）c=b（b（bc））。很清楚，这可同样好地适用于任何其他自然数。（事实上λabc.[（ab）（bc）]可以和S一样好地做到。）

把一个数乘二又如何呢？这种加倍可由

D=λabc.[（ab）（（ab）c）]

获得，它可再次由作用于3上而得到验证：

$$D=\lambda abc.[(ab)((ab)c)]\mathbf{3}=\lambda bc.[(\mathbf{3}b)((\mathbf{3}b)c)]=\lambda bc.[(\mathbf{3}b)(b(b(bc)))]=\lambda bc.[b(b(b(b(bc))))]=\mathbf{6}。$$

事实上，加法、乘法和自乘的基本算术运算可分别定义为：

A=λfgxy.[（（fx）（gx））y]，

$M=\lambda fgx.[f（gx）]$，

$P=\lambda fg.[fg]$。

实际上，读者可以确信——要不就不加考察信以为真，即：

$$(\mathbf{A}\mathbf{m})\mathbf{n}=\mathbf{m}+\mathbf{n},\ (\mathbf{M}\mathbf{m})\mathbf{n}=\mathbf{m}\times\mathbf{n},\ (\mathbf{P}\mathbf{m})\mathbf{n}=\mathbf{n}^{\mathbf{m}},$$

其中m和是n丘奇的两个自然数的函数，m+n是它们和的相应函数，等等。最后那个公式是最令人惊异的。让我们仅仅验证其m=2，n=3的情形：

$$(\mathbf{P}\mathbf{2})\mathbf{3}=((\lambda fg.[fg])\mathbf{2})\mathbf{3}=(\lambda g.[\mathbf{2}g])\mathbf{3}=$$

$$(\lambda g.[\lambda fx.[f(fx)]g])\mathbf{3}=$$

$$\lambda gx.[g(gx)]\mathbf{3}=\lambda x[\mathbf{3}(\mathbf{3}x)]=$$

$$\lambda x.[\lambda fy.[f(f(fy))](\mathbf{3}x)]=$$

$$\lambda xy.[(\mathbf{3}x)((\mathbf{3}x)((\mathbf{3}x)y))]=$$

$$\lambda xy.[(\mathbf{3}x)((\mathbf{3}x)(x(x(xy))))]=$$

$$\lambda xy.[(\mathbf{3}x)(x(x(x(x(x(xy))))))]=$$

$$\lambda xy.[x(x(x(x(x(x(x(x(xy)))))))))]=$$

$$\mathbf{9}=\mathbf{3}^{2}$$

减法和除法不是这么容易定义的（我们的确需要某种当m比n小

时“m-n”以及当m不能被n整除时“m÷n”的约定。事实上，20世纪30年代早期，克林发现如何在丘奇的方案中表达减法运算就被认为是这一学科的重要里程碑！后来接着又有其他的运算。最后，丘奇和图灵在1937年独立地指出，不管什么样的可计算的（或算法的）运算（现在在图灵机的意义上）都可以按照丘奇的一种表达式获得（而且反之亦然）。

这是一个真正惊人的事实，它被用来强调可计算性思想的基本客观性以及数学特征。初看起来，丘奇的可计算性概念和计算机器没有什么关系。然而，它和实际计算具有某些基本关系。尤其是，有力而灵活的电脑LISP语言以一种根本的方式参与到丘奇计算法的基本结构中来。

正如我早先指出的，还有其他定义可计算性概念的方法。波斯特的计算机器的概念和图灵的非常接近，并且几乎是同时独立提出的。近世还有更有用的可计算性（递归性）的定义，这是J.赫伯拉德和哥德尔提出的。H.B.邱雷在1929年，以及M.轩芬克勒在1924年稍早些时候提出了不同的方法，丘奇计算法就是部分地由此发展而来（参见*Gandy 1988*）。现在研究可计算性的手段（诸如在*Cutland 1980* 中描述的一台无限记录机器）在细节上和图灵原先的相差甚多，而且它们更实用得多。然而，不管采用那种不同的手段，可计算性的概念仍然相同。

正如许多其他的数学观念，尤其是更漂亮的、更基本的那些，可计算性的观念似乎自身具有某种柏拉图的实在性。在下面两章，我们应该探讨的正是数学概念的柏拉图实在性的这个神秘问题。

第3章　数学和实在

托伯列南国

想象我们到某一遥远世界作远程旅行。我们称这一遥远世界为托伯列南国。现在把我们的遥感仪器收集到的信息展现在面前的屏幕上。调好焦距后就看到了图3.1。

它为何物？它是一只形状古怪的昆虫吗？也许它是一个深颜色的并有许多山溪注入的湖泊。也许它是一座巨大的形状奇特的异国城市，公路沿着不同方向散开到附近的小镇和乡村去。它也许为一个岛屿——让我们寻找看在附近是否有和它相连接的陆地。我们可以后退一些，把我们感觉仪器的放大倍数减少到原来的1/15左右。嗬，整个世界进入了我们的视界之内（图3.2）。

我们的“岛”在图3.2中看起来成为标记“图3.1”下的小斑点。除了一条连接到右手的裂缝上去的以外，从原先岛上出发的小片断（溪流、路径、桥梁？）全部都终结了。该裂缝最终接到我们在图3.2画出的大得多的物体上去。这个更大的物体虽然和我们第一次看到的岛不完全一样，但明显地相似。如果我们更仔细地审视这一物体和海岸线相像的东西，就发现多得数不清的圆形的瘤状结构。每一结构自身又具有类似的瘤。似乎每一小瘤都在某一微小的地方附在一个更大的瘤上，由此在大瘤上产生出许许多多的小瘤。当图像变得更清楚时，人们就看到了从这个结构发出的成千上万根的细丝。这些细丝在不同的地方分叉并常常剧烈地弯折。在细丝的某些点，我们似乎看到了具有现有的放大倍数的感觉仪器所不能分析的复杂纽结。很显然，这物体不是实际的岛屿或陆地，也不是任何风景。或许我们看到了某种怪诞的甲虫。我们首先看到的是它的婴儿，它用某种丝线状的脐带安静地把自己连接在母体上面。

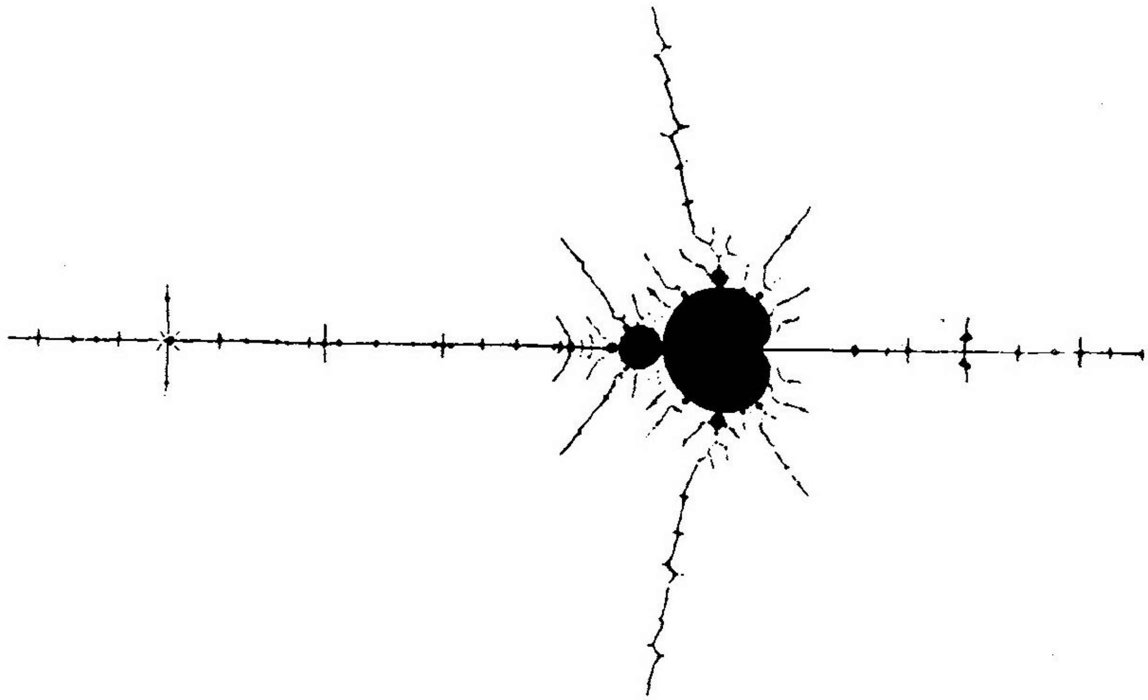


图3.1　奇异世界之第一瞥

让我们把感觉仪器的放大倍数提高10倍，再来考察这个怪物的一个瘤的性质（图3.3—其位置在图3.2中的“图3.3”的标志的下面）。这个瘤本身和怪物整体非常相似——除了在接触点以外。请注意在图3.3中的不同地方5根细丝并到一块。这个特定的瘤似乎有一确定的“五性”（正如在最上面的瘤具有“三性”一样）。如果我们考察下一个相当尺度的瘤，在图3.2中稍微向左下方一点，我们就会在附近发现“七性”，再下一点为“九性”，并以此类推。当我们进入图3.2中的两个最大区域之间的裂缝，就会发现右边的瘤以奇数来表征，每回增加2。让我们钻到裂缝深处，把图3.2再放大10倍左右（图3.4）。我们看到其他许多小瘤以及扭转的结构。在右边称为“海马谷”的区域可鉴别出某些微小的涡旋状的“海马尾巴”——如果放大倍数足够大的话，我们就将看到不同的“海乌贼”或者别具花样的区域。这也许的确是某种奇异的海岸线——也许是所有各色各样生命产生的珊瑚。看起来像是花的东西在更高的放大倍数下显得是由成千上万个微小，但同时却是不可思议的复杂的结构组成，每一结构都有极多的丝状物和扭转的涡旋尾巴。让我们稍微仔细地考察一个较大的海马的尾巴，也就是在图3.4中刚好能见到标志为“图3.5”的那个（它附在具有“29性”的瘤上面！）。大约再放大250倍左右，我们就得到了画在图3.5中的涡旋。我们发现这个尾巴非同寻常，它是

由最复杂的、前后扭曲的、无数的小涡旋以及像章鱼和海马那样的区域组成。

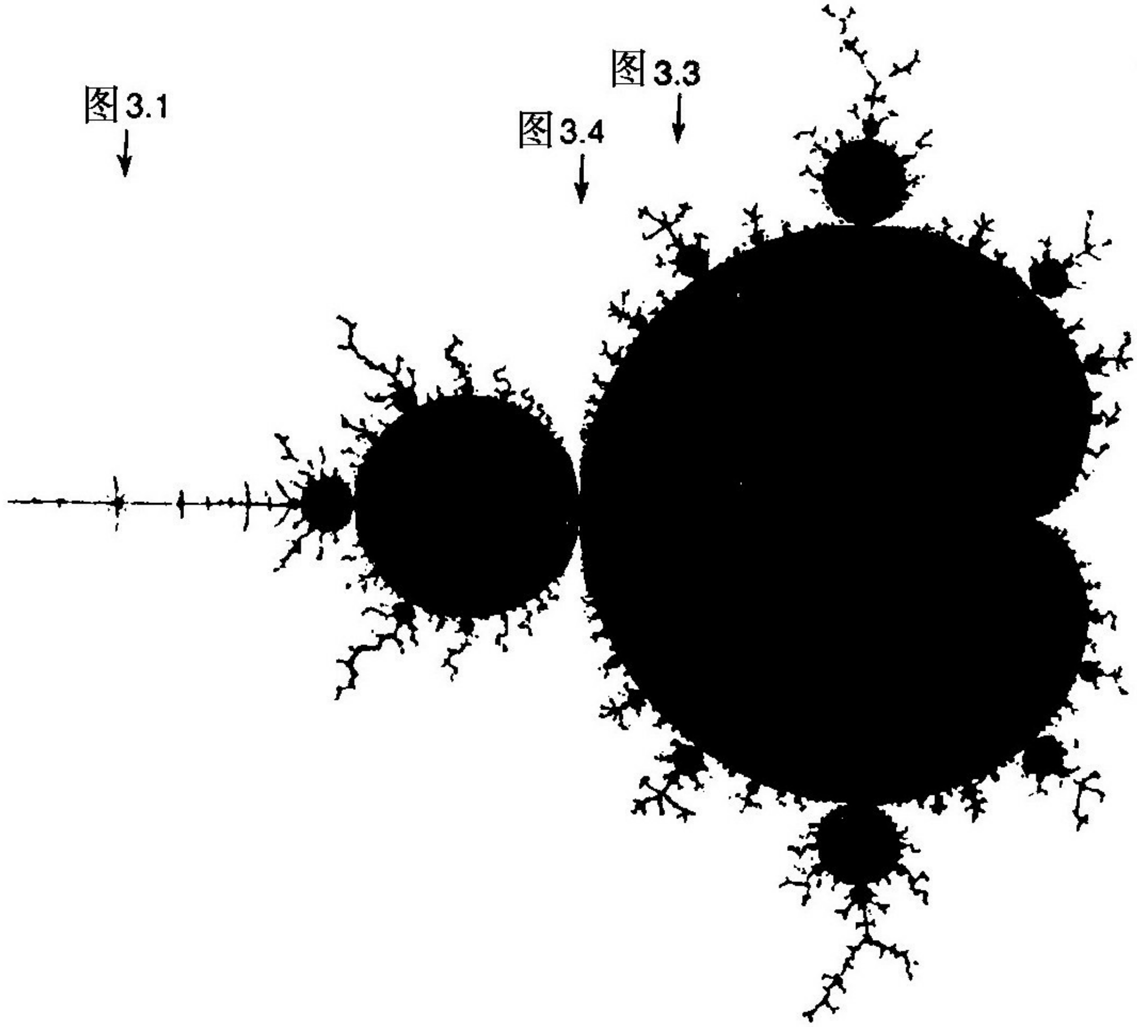


图3.2　整个“托伯列南国”。箭头之下标出了在图3.1、图3.3和图3.4中的放大部分的位置

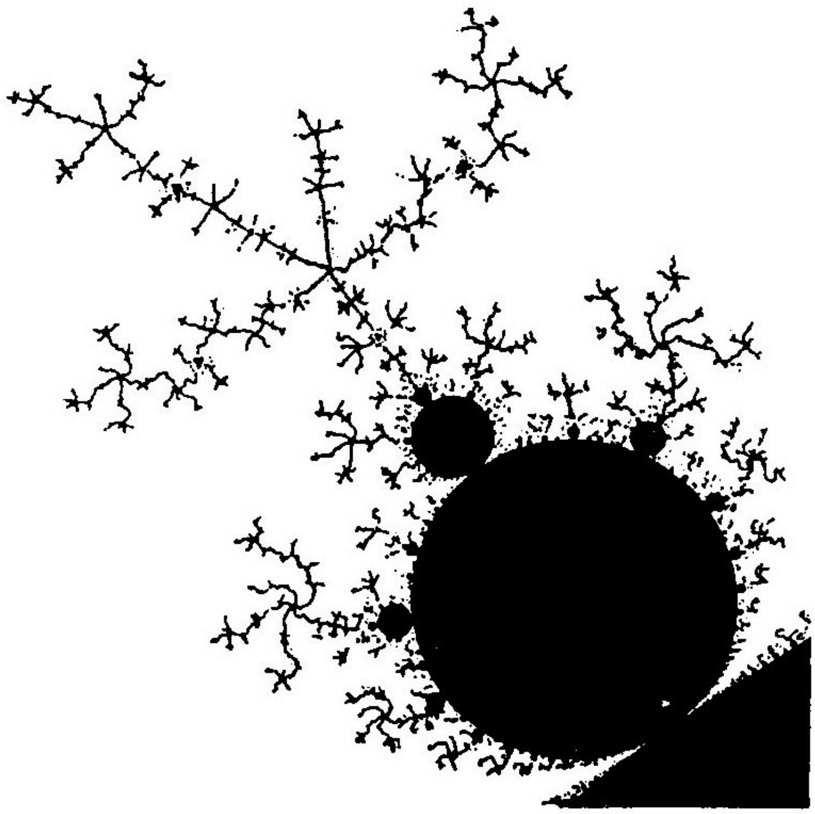


图3.3　一个具有“五性”的细丝的瘤

在这个结构的许多地方刚好有两个涡旋碰到一起。让我们把放大倍数增加30倍左右，以考察其中一处（在图3.5中的标志“图3.6”的下面）。请注意，我们是否发现了中间有个奇怪但非常熟悉的对象？再放大6倍左右（图3.7）就能揭示出一个怪物的小婴儿——它几乎和我们考察过的整个结构完全一样！如果我们细看，就会发现从它那里出发的细丝和从主结构那里出来的略有差别。它们扭曲并延伸到更远得多的距离去。然而比细小结构本身几乎和它的上一代毫无差别，甚至在非常相应的地方拥有自己的后代。如果我们还进一步放大，就能继续考察这些东西。孙子们又非常类似于它们的共同祖先——人们很容易相信，这些现象会无限地延续下去。只要不断地提高我们感觉仪器的放大倍数，就可随心所欲地探索托伯列南的奇异世界。我们发现了无穷尽的变化：没有两个区域是完全相像的——但是我们很快就会习惯于存在的一些普遍的风格。而熟知的类甲虫的结构以越来越小的尺度重新出现。每一回它的附近的细丝结构都和早先看到的不同，并以不可置信的复杂的美妙的新景象呈现在我们的面前。

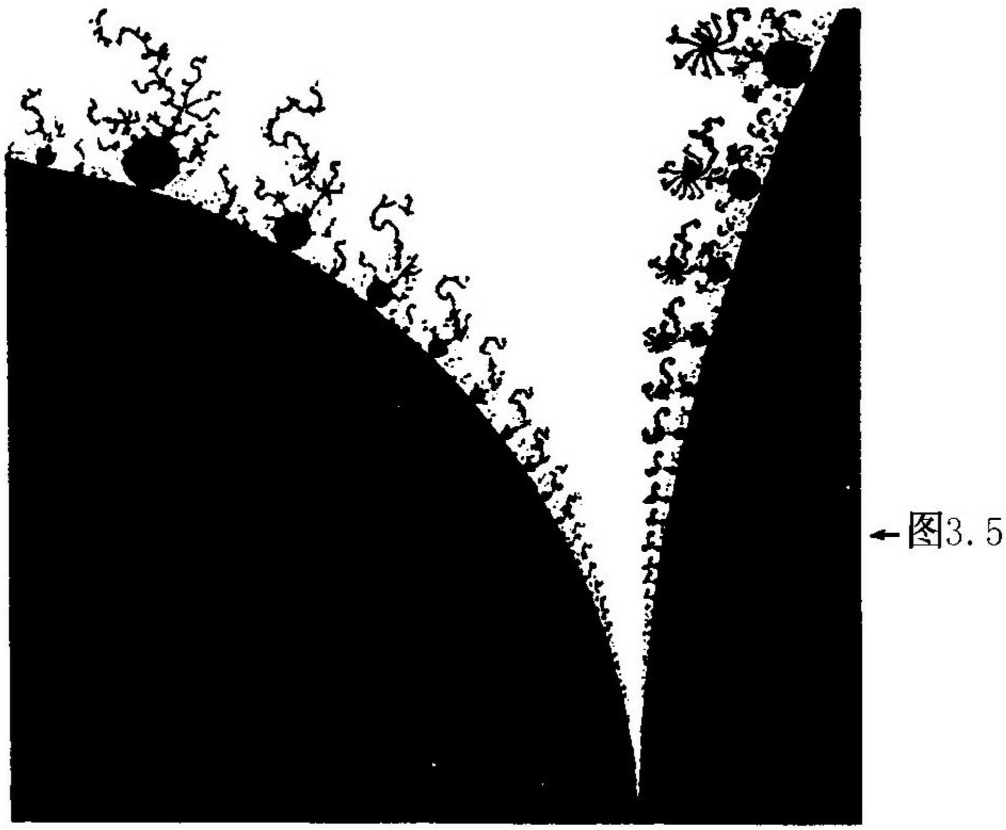


图3.4　主狭缝：在右下方可见到“海马谷”

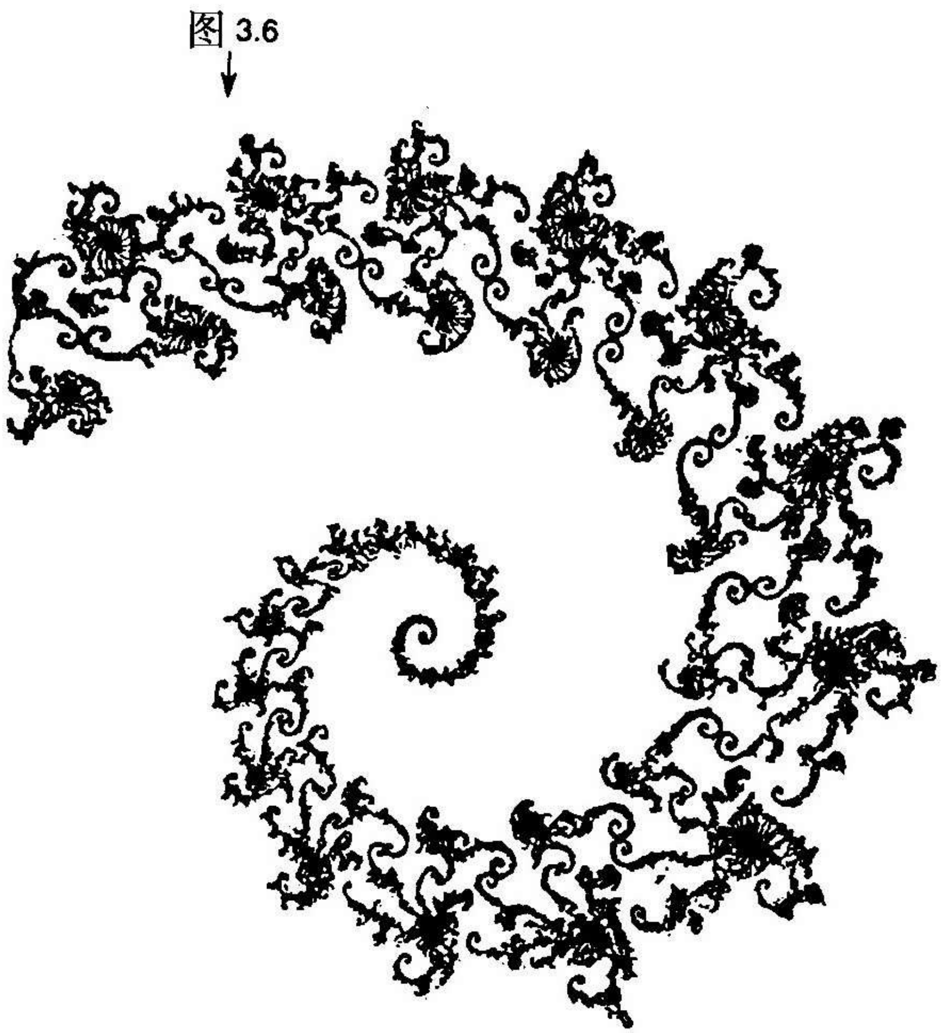


图3.5　海马尾巴的近窥

使我们目瞪口呆地奇异的、变化多端的、美妙的、复杂的国土究竟为何物呢？许多读者无疑已经知道。但还有一些读者不知道。这世界只不过是一点抽象数学——称为芒德布罗集[1]的集合。尽管它无疑是复杂的，却是由极其简单的规则产生的！为了恰当地解释该规则，我首先得解释什么是复数。除了这里以外，在将来还有用。它对于量子力学的结构，所以也就是我们生活其中的世界的运行是绝对基本的。它们构成了数学中的一个伟大奇迹。为了解释何为复数，首先得提醒读者何为“实数”。另外，弄清概念和“真实世界”的实在的关系也是非常有益的。



图3.6　两个涡旋会合处的进一步放大细节。在中心点处刚刚可以见到一个小婴儿

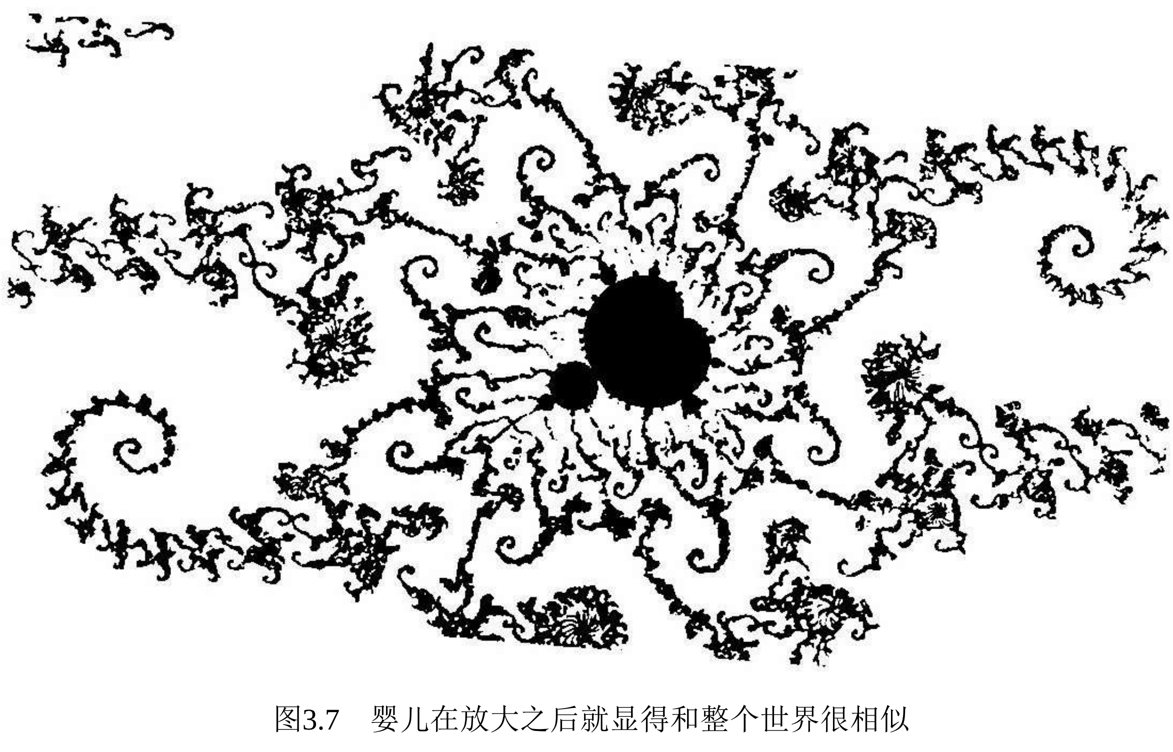


图3.7　婴儿在放大之后就显得和整个世界很相似

实数

我们知道自然数可被罗列如下：

0，1，2，3，4，5，6，7，8，9，10，11，……，

这些是不同种类数中最初等和最基本的。任何分立的对象都可以用自然数予以量化：我们可以讲田地里有27只绵羊，可以讲2次闪电，12个晚上，1000个词，4次谈话，0个新观念，1个错误，6位缺席者，2次方向改变等。自然数可以相加或相乘以得到新的自然数。它便是上一章给出的关于算法的一般讨论的对象。

然而某些重要的运算会把我们带到自然数王国之外——最简单的是减法。为了系统地定义减法，我们需要负数；为此目的我们建立了整数的整个系统

……，-6，-5，-4，-3，-2，-1，0，1，2，3，4，5，6，7，…….

某些事物，譬如电荷、银行的存款或者年份[13]可用这类数来量化。然而，这些数的范围仍然过于局限。这是由于把一个数除以另一个数时，我们仍然不能畅通无阻。相应地，我们需要分数或有理数。

0，1，-1，1/2，-1/2，2，-2，3/2，-3/2，1/3，…….

这一些对于有限算术的运算已经足够。但是为了许多更好的目的，我们还得走得更远些，以包括无穷或极限运算。例如，大家熟悉的在数学上极其重要的量π就以多个这类无穷式出现。特别地，我们有：

π=2{（2/1）（2/3）（4/3）（4/5）（6/5）（6/7）（8/7）（8/9）……},

以及

π=4（1-1/3+1/5-1/7+1/9-1/11+……）。

这些都是著名的表式。第一式是由英国数学家、语法学家兼速算家约翰·沃利斯在1655年首次得到的；而第二式实际上是苏格兰数学家兼

天文学家（以及第一台反射望远镜的发明者）詹姆斯·格里高里在1671年得到的。正如π那样，以这种方法定义的数不必是有理数（也就是不具有n/m的形式，这里n和m是整数，m不为零）。为了包括这样的量，数的系统必须被推广。

这个推广的数的系统被称为“实”数系统——就是那些可以无尽小数展开的熟悉的数，譬如：

-583. 70264439121009538…….

按照这样的表述，π可写成众所周知的表式：

π=3.14159265358979323846……。

还能以这种方法表达的数种，有正有理数的平方根（或立方根或四次方根等），例如：

2=1. 41421356237309504……;

甚或任何正实数的平方根（或立方根等），正如伟大的瑞士数学家列纳多·欧拉发现的π的表示：

$$\pi = \sqrt{\left\{6(1+1/4+1/9+1/16+1/25+1/36+\cdots\right\}} \text{。}$$

实数实际上是我们日常必须打交道的数种，虽然通常我们仅仅关心它们的近似值，只要展开到很少的几位小数位就满意了。然而，在数学的陈述中我们要准确地指定实数，要求某种无穷的诸如整个无穷小数展开的描述，或者也许如上述由沃利斯、格里高里和欧拉给出的π的其他的无穷的数学表达式。（我将通常用小数展开，只是因为这些是最熟悉的。对于数学家而言，存在不同的令人更满意的表达实数的办法，但我们在这里不必为之操心。）

人们也许会感到处理全部的无尽展开是不可能的。但事情并非如此。简单的反例是：

1/3=0. 333333333333333……

这儿的点表明3的序列将无尽地延伸下去。为了处理这个展开，我们所需要知道的是，只要肯定这个展开以同样的3的方式无限地继续下去就行了。任何有理数都有重复（或有限）的小数展开，例如：

93/74=1. 2567567567567567……

此处序列567无限地重复下去，而这可以被完全地处理。而表式：

0.220002222000000222222000000002222222220　……

定义一个无理数，也一定可以被完全处理（每一次0序列和2序列都增加一位）。还能给出许多熟知的例子。在每一种情形下，只要我们知道展开所根据的法则也就满意了。如果有某种产生连续位数的算法，则该算法就提供我们处理整个无尽小数展开的方法。其展开可被算法产生的实数称为可计算数（在这里使用十进制，而不用譬如讲二进制展开，并没有什么深意）。刚才考虑的π和2是可计算数的例子。在每一种情况下，仔细叙述这些规则是稍微有些复杂，但在原则上并不难。

然而，在这个意义上还有许多不可计算的实数。我们在上一章已经看到，存在不可计算的但仍为完好定义的序列。例如，我们可取一个小数展开，其n位数取1或取0依图灵机作用到n时停止或不停止而定。一般地讲，对于一个实数，我们仅仅要求必须有某种无尽的小数展开。我们不要求是否有一产生第n位数的算法。我们甚至也不要知道在原则上实际定义该n位数的规则[2]。可计算数是很难纠缠的东西。即使只处理可计算数，人们也不能够使它的所有运算保持为可计算的。例如，甚至一般地去决定两个可计算数是否相等也不是可计算的事体！由于这类原因，我们宁愿处理所有的实数。在这里小数展开可以是任意的，而不必只是可计算序列等。

最后，我倒是要指出，在结尾以无穷个接续的9和无穷个接续的0展开的实数之间有一等同；例如：

-27. 1860999999……=-27.1861000000……

有多少个实数

让我们喘息一下，来鉴赏在从有理数过渡到实数时所得到的推广的广阔性。最初人们也许会以为，整数的个数比自然数的更多，由于每一自然数都是整数，而某些整数（也就是负的）不是自然数。类似地，人们也会以为分数的数目比整数的数目更多。然而事情并非如此。按照极有创见的俄裔德国数学家——乔治·康托尔在19世纪后半叶提出的强有力的美丽的无限数理论，分数的总数目、整数的总数目和自然数的总数目是同一无穷数，均用$\aleph_0$（“阿列夫零”）来表示。（值得注意的是，在大约250年前的17世纪初叶，伟大的意大利物理学家和天文学家伽利雷·伽利略也部分地预料到这一类思想。在第5章将会提到伽利略的其他一些成就。）人们可用如下建立的“一一对应”的办法来显示整数和自然数具有同样的数目：

整数		自然数
0	↔	0
−1	↔	1
1	↔	2
−2	↔	3
2	↔	4
−3	↔	5
3	↔	6
−4	↔	7
⋮	⋮	⋮
−n	↔	2n−1
n	↔	2n
⋮	⋮	⋮

请注意，每一整数（在左列）和每一自然数（在右列）在表中出现一次并只有一次。在康托尔的理论中像这样的一一对应的存在建立了左列物体的数目和右列的是一样的命题。这样，整数的数目的确和自然数的数目一样。在这种情形下数目为无穷，但这没关系。（发生在无穷数中的仅有的古怪事情是，我们可以从一个表上取走一些数而仍然能找到两个表之间的一一对应！）以某种类似的但更复杂的形式，人们可在分数和整数之间建立起一一对应[为此我们可以采用把一对自然数（分子和分母）代表为一个单独自然数的方法]。可以和自然数建立一一对应关系的集称为可数的。所以，可数的无限集共有$\aleph_0$ 个元素。我们现在看到了，整数是可数的，所有的分数也是如此。

有没有不可数的集合呢？虽然我们进行了自然数首先到整数、然后到有理数的推广，但是我们实际上并没有增加所处理对象的总数。也许读者已得到印象，以为所有无穷集都是可数的。不对，在推广到实数时情况就变得非常不同。康托尔的一个最重大的成就是，他指出了，在实际上实数比有理数有更多的数目。康托尔进行论证的办法在第2章被称为“对角线方法”。这个方法被图灵用来表示图灵机的停机问题是不可解的。康托尔的论证，正如图灵的办法，是用反证法的步骤。假定我们所要建立的结果是错误的，也就是所有的实数的集是可数的。那么在0和1之间的实数肯定为可数的，而我们存在某种列表，可将实数和自然数之间进行一一配对，譬如：

自然数		实数
0	↔	0.**1**0357627183…
1	↔	0.1**4**329806115…
2	↔	0.02**1**66095213…
3	↔	0.430**0**5357779…
4	↔	0.9255**0**489101…
5	↔	0.59210**3**43297…
6	↔	0.636679**1**0457…
7	↔	0.8705007**4**193…
8	↔	0.04311737**8**04…
9	↔	0.786350811**5**0…
10	↔	0.4091673889**1**…
⋮	⋮	⋮

我已把对角线上的数字用黑体字写出。对于这一特殊的表，这些数字分别为：

1，4，1，0，0，3，1，4，8，5，1，……

而对角线方法步骤是（在0和1之间）构造一个实数，其小数展开（在小数点后）在每一对应的位数上和这些数字都不同。为了确定起见，让我们讲，只要对角线数和1不同的都为1，而对角线数为1的都为

2。我们在现在情况下就得到了

0.21211121112 ……

的实数。这个实数不可能出现在我们的表上。这是因为它在（小数点后的）第一个小数位上和第一个数不同，在第二个小数位上和第二个数不同，在第三个小数位上和第三个数不同等。由于我们假定这个表包含所有在0和1之间的实数，所以这是一个矛盾。这一矛盾导致我们所要证明的，也就是说，在实数和自然数之间没有一一对应。相应地，实数的数目实际上比有理数的数目更大，因而不是可数的。实数的数目是标有C的无限数。（C的意思是连续统，这是实数系统的另一名字。）人们会问，譬如讲，为何这一个数目不叫作ℵ1呢？事实上符号$\aleph_1$ 是用来表示比$\aleph_0$ 大的下一个无限数。去决定事实上$C=\aleph_1$ 成立与否是一道著名的被称为连续统假设的未解决问题。

顺便可以提及，可计算数是可数的。为了数这些数，我们只要按照数字的顺序列出那些产生实数的图灵机（也就是产生实数连续数字的机器）。我们可望从这表中除去产生任何早先出现在表中的实数的图灵机。由于图灵机是可数的，所以可计算的实数也一定是可数的。我们为何不能把对角线方法应用到该表上以产生一个不在该表的新的可计算数呢？回答是基于这样的一个事实，即我们不能一般地可计算地确定，一台图灵机是否在这表上。为了做到这一点，事实上也就涉及我们能够解决停机问题。有的图灵机，可以开始产生一个实数的数字，然而停住而永远不再产生另一数字（因为它“不停机”）。没有可计算的方法去决定哪一台图灵机会以这种方法卡住。这基本上是停机问题。这样，我们对角步骤会产生某实数，这数不是可计算的。这个论证事实上可用于表明不可计算数的存在。图灵用于显示不能算法地解决的，正如在上一章所罗列的各类问题的存在，正是精确地沿用了这种推理方法。我们在后面还会看到对角线方法的其他应用。

实数的“实在性”

我们先不管可计算性的概念。由于实数似乎提供了测量距离、角度、时间、能量、温度或者许多其他几何和物理量的大小，所以被叫作“实”的。然而在抽象定义的“实”数和物理量之间的关系，不像人们所

想象的那么一目了然。实数点被当成数学的理想化，而不是任何实际物理客观的量。例如，实数系统具有如下性质，在任何两个实数之间必有另一个实数，而不管该两数靠得多近。人们根本就不清楚，物理的距离或时间是否现实上具有这一性质。如果我们不断地对分两点之间的物理距离，最后就会到达这样微小的尺度，以至于在通常意义下的距离概念本身不再具有意义。人们预料在亚原子粒子的$1/10^{20}$ 的“量子引力”尺度下 [14] ，这的确会发生。但是为了和实数相匹配，我们就必须走到比它小得任意多的尺度：例如$1/10^{200}$ ，$1/10^{2000}$ 或$1/10^{20000}$ 的粒子尺度。人们一点也不清楚，这么荒谬的微小尺度究竟有什么物理意义。类似的议论也适用于相应的微小的时间间隔。

物理学选用实数系统的原因在于它的数学上的可用、简单、精巧以及在非常广大的范围内和距离以及时间的概念相符合。它之所以被选用并不是因为知道它和这些物理概念在所有的范围中都一致。人们还可以预料到，在非常微小的距离或时间的尺度下，不存在这样的一致。人们通常用尺来测量简单的距离，但这样的尺在我们追溯到它们自身原子的尺度时，就变得粗糙起来。这一切并不妨碍我们继续准确地利用实数，但要经过更加精细的处理，才能测量更小的距离。我们至少要有点怀疑，在极小尺度的距离下，也许最终存在有根本原则上的困难。自然对于我们真是恩惠有加，我们从小习惯用于描述日常或更大尺度的事物的同一实数，在尺度比原子小很多，肯定在比“经典”的亚原子粒子，譬如电子或质子的经典直径小百倍的尺度下仍然有用，似乎直到比这粒子小20个数量级的“量子引力尺度”仍然适用。从经验得知，这是极不寻常的推论。熟知的实数距离的概念似乎还可外推到最遥远的类星体以及更远处，给出了至少10^{42} 。也许10^{60} 甚至更广的大范围。事实上，实数系统的适当性通常是不可置疑的。我们原先和实数相关的经验主要被限于相对有限的范围，人们为什么对实数于物理精密描述的可用性如此信心百倍呢？

这种信念——也许是不当的——必须来源于（虽然这个事实经常不被承认）实数系统逻辑的优雅、一致性和数学的威力以及对自然的深刻数学和谐的信仰。

复数

实数系统并没有全揽数学的威力和优雅品格。其中仍有一些讨厌之处，例如只能对正数（或零）而不能对负数取平方根。先不讲和物理世界有直接关系的任何问题，单从数学的观点知道，如果能像处理正数那样对负数求平方根，那就极其方便了。让我们简单地假定，或“发明”数-1的平方根。我们用i来表示它，所以就有：

$i^2=-1$

当然i的数量不能是实数，因为任何实数自乘的结果总是正数（或是零，零自乘得零）。由于这个原因，习惯上用“虚数”来称呼其平方为负数的数。正如我早先强调的，“实”数和物理实在的关系不像初看起来那么直接、那么令人信服，这里实际牵涉到数学的无限精细化的理想化，自然并没有先天地保证这种做法的合理性。

一旦有了-1的平方根，就可以不费劲地得到所有实数的平方根，如果a为一个正实数，则量

$i\times\sqrt{a}$

是负实数-a的平方根。（还有另一平方根-i×a。）i本身又如何呢？它有平方根吗？它的确有，很容易检验量

（1+i）/$\sqrt{2}$

（以及其负量）的平方得i。这个数有平方根吗？答案又是肯定的；量 $\sqrt{\frac{(1+1\sqrt{2})}{2}}+i\sqrt{\frac{(1-1\sqrt{2})}{2}}$ 或者它的负量的平方的确为（1+i）/$\sqrt{2}$ 。

我们注意到，在形成这样的量时，我们允许把实数和虚数相加，也允许把我们的数乘任意实数（或除以非零的实数，这相当于乘以它们的倒数）。所得的结果称为复数。复数是具有形式

$a+ib$

的数，这里a和b是实数，分别称作该复数的实部和虚部。将这样的两个数相加和相乘必须遵循通常的代数法则以及$i^2=-1$的规则：

（a+ib）+（c+id）=（a+c）+i（b+d），

（a+ib）×（c+id）=（ac-bd）+i（ad+bc）.

现在出现值得注意的情况！我们对这个系统的动机是使对任何数都能取平方根。这个目的是达到了，虽然还不这么明显。但是，它做得比这还多得多：取立方根、5次方根、99次方根、π次根、（1+i）次根等都可以畅通无阻地进行（正如伟大的18世纪数学家列纳多·欧拉指出的那样）。作为复数的另外一个魔术，我们考察在中学就学到的三角几何中略显复杂的公式，两个角之和的正弦与余弦公式

sin（A+B）=sinAcosB+cosAsinB，

cos（A+B）=cosAcosB-sinAsinB，

只不过分别是简单得多（也容易记忆得多！）的复方程[15]

$$e^{iA+iB}=e^{iA}e^{iB}$$

的虚部和实部。

我们在这里所要知道的是“欧拉公式”（众所周知，在欧拉之前多年，该公式就被16世纪杰出的英国数学家罗杰·可提斯得到）

e^{iA} =cosA+isinA，

把它代入前面的方程，其结果的表达式为

cos（A+B）+isin（A+B）=（cosA+isinA）（cosB+isinB），

只要把右边乘出，我们就得到所需的三角关系式。

而尤其值得注意的是，任何代数方程

$$a_0+a_1z+a_2z^2+a_3z^3+\cdots+a_nz^n=0$$

（此处a_0，a_1，a_2，……，a_n为复数，$a_n \neq 0$）总有复数解。

例如，存在满足关系

$$z^{102}+999z^{33}-\pi z^2=-417+i$$

的一个复数z，虽然这一点绝非明显！这一个普遍的事实有时被称作“代数基本定理”。不少18世纪的数学家都为证明这个结果奋斗过。甚至欧拉也没有找到一个满意的一般的论证。后来在1831年，伟大的数学家和科学家卡尔·弗列得里希·高斯给出了惊人的独创的论证，并提供了第一个一般性证明。他的证明的关键部分是几何地表达复数，然而利用拓扑学[16]的论断。

高斯实际上不是使用复数描述的第一个人。沃利斯在大约200年前就这么做了，虽然他没有像高斯那样有效地使用这工具。通常把复数的几何表示归功于瑞士的簿记员金·罗伯特·阿伽德，他在1806年将其描述出来，尽管挪威的测绘家卡斯帕·温塞尔事实上在9年前就给出了完整的描述。为了和这个惯用的（虽然与历史不符）术语相一致，我将复数的标准几何表示称为亚根平面。

亚根平面是一个通常的欧几里得平面，它具有标准笛卡儿的x，y坐标，x标出水平距离（向右为正，向左为负），而y标出垂直距离（向上为正，向下为负）。复数

$z=x+iy$

在亚根平面中以坐标为

（x, y）

的点所表示（图3.8）。

注意0（作为一个复数）由坐标系的原点代表，1是由x轴上的特殊的点代表。

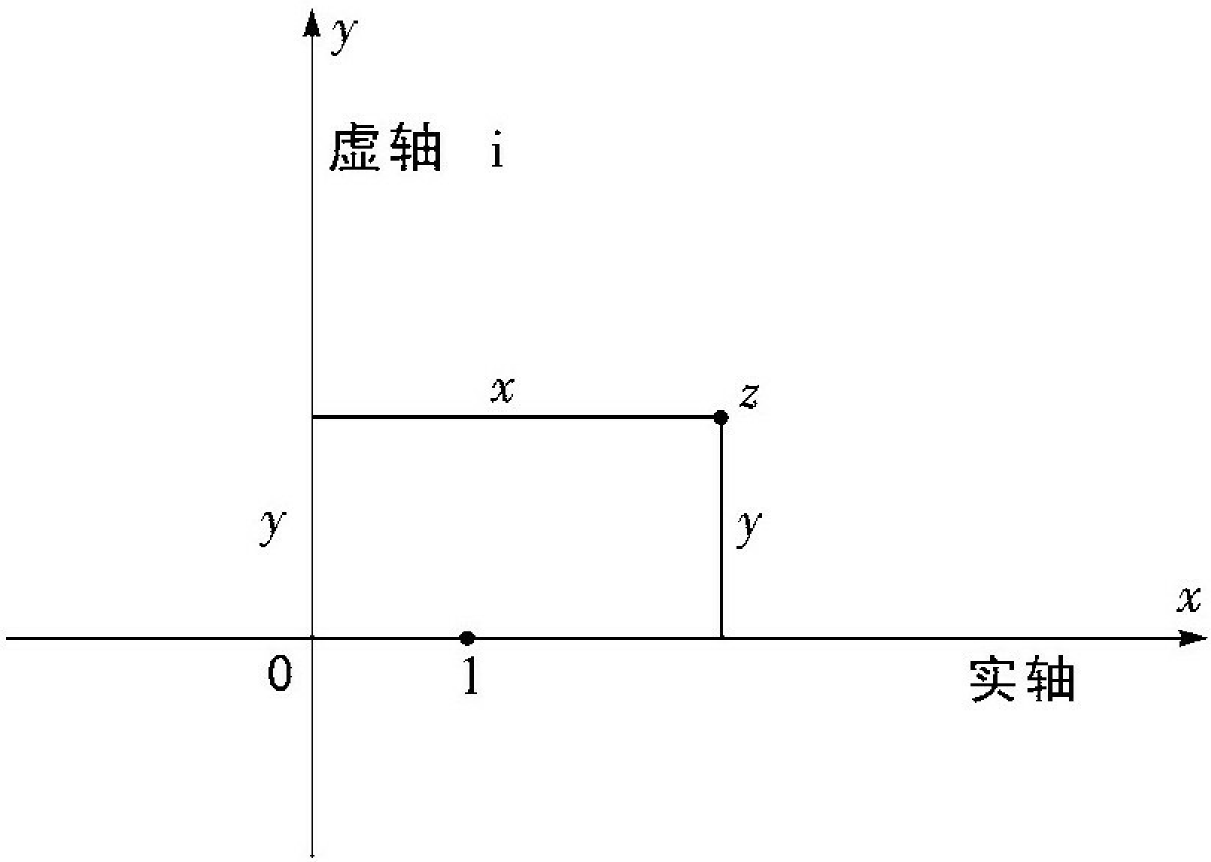


图3.8　在亚根平面上画出了复数z=x+iy

亚根平面为我们把整个复数的家族组织成一个几何上有用的图形。这类事对我们而言并无新奇之处。我们已经熟悉实数可以组织成为一个几何的图像的方法，也就是一根向两个方向无限延伸的直线。直线上的特定点为0，另一点为1。点2的位置处于它到1的位移和1到0的位移相同的地方；点1/2处于0和1的中点；点-1使得0处于它和1的中间，等等。以这种方式标出实数的集合称为实线。对于复数，我们事实上用两个实数作为复数a+ib的坐标，也就是a和b。这两个数给出我们一个平面——亚根平面上的点的坐标。例如，我在图3.9上近似地标出了复数的位置。

u=1+i1. 3，v=-2+i, w=-1.5-i0.4

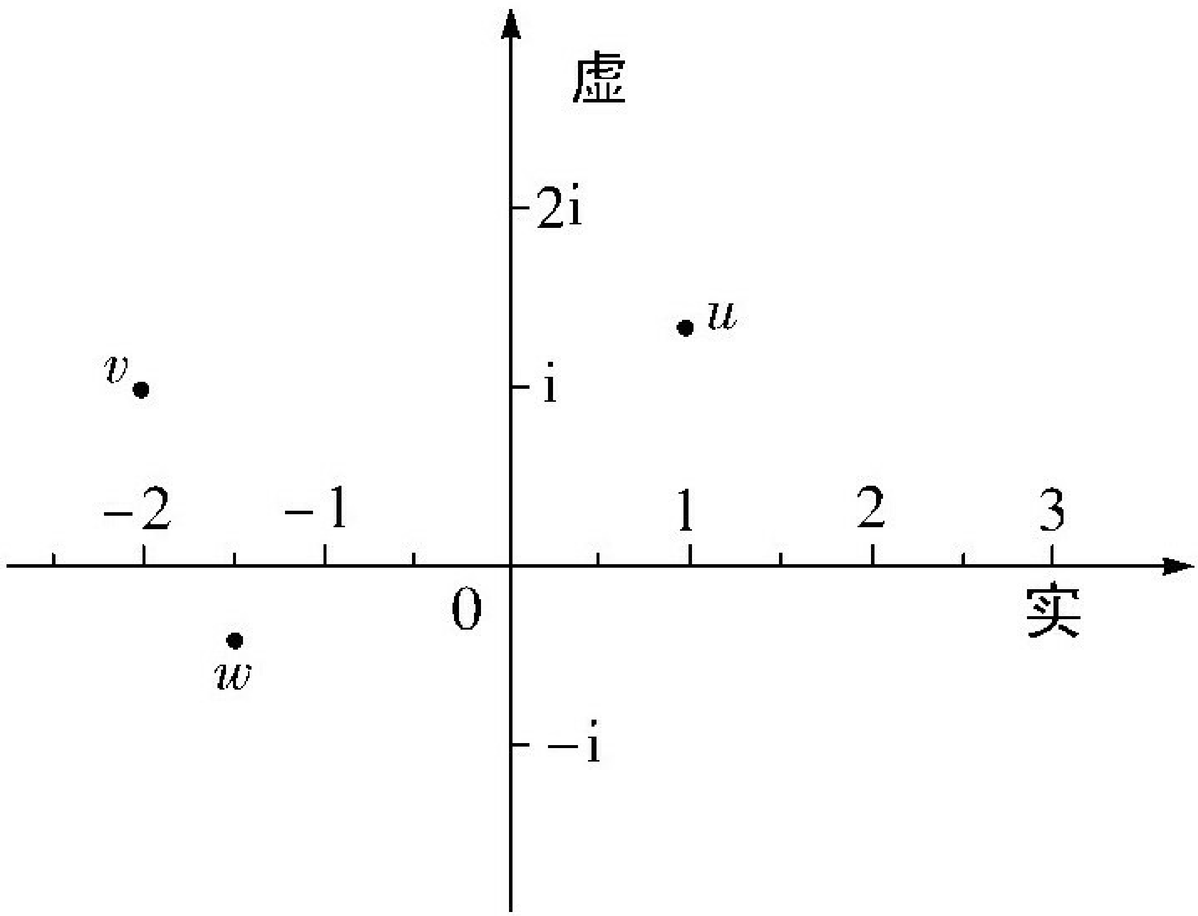


图3.9　亚根平面上的u=1+i1.3，v=-2+i和w=-1.5-i0.4的位置

现在复数的加法和乘法的基本代数运算具有清楚的几何意义。首先考虑加法。假设u和v为两个复数，并按照上述的方案表示在亚根平面上。则它们的和u+v就由这两点的“向量和”来表示；也就是说，它处于由u, v和原点0构成的平行四边形的另一顶点。我们不难看出，由这种构造（图3.10）可以得到和，但是我在这里把证明省略掉。

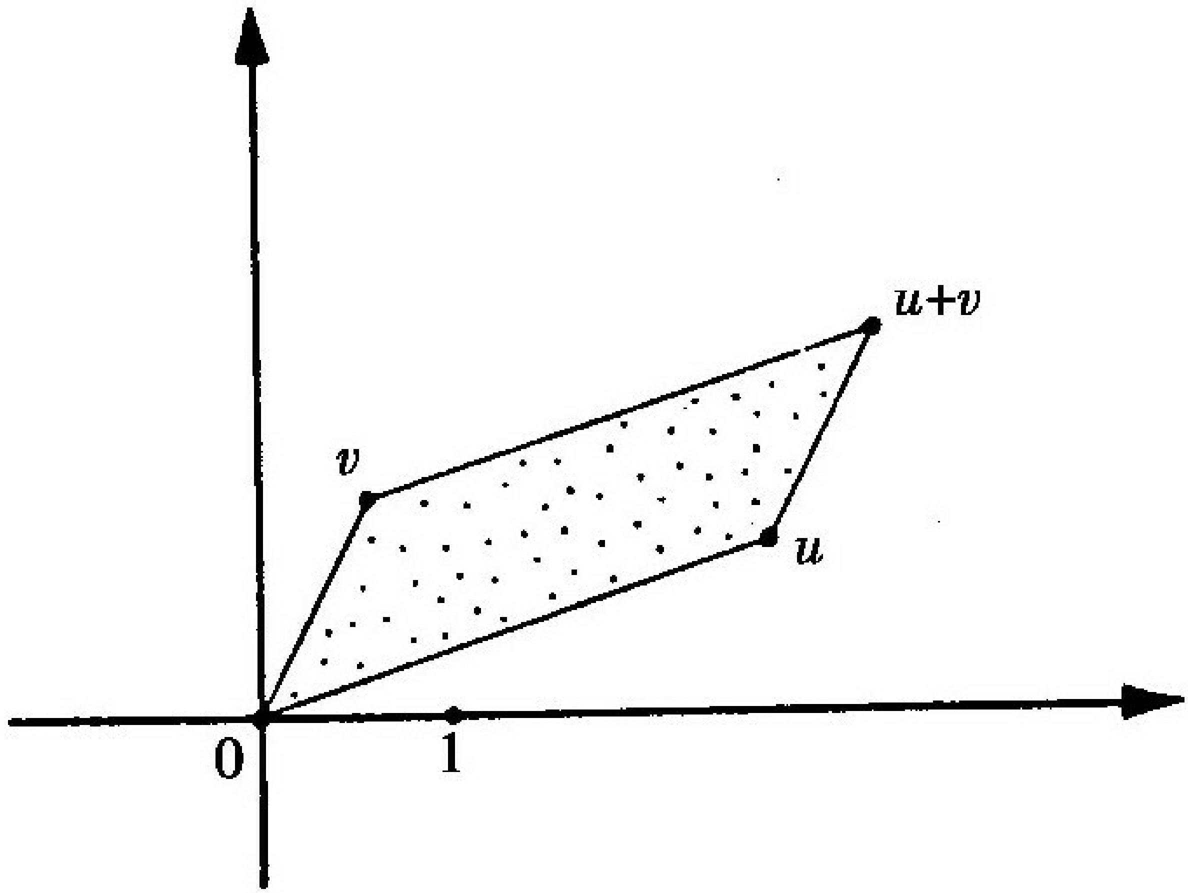


图3.10　两个复数u和v的和u+v可由平行四边形定律得到

乘积uv也有清楚的几何解释（图3.11），这稍微不太容易看得出来。（我在这里又省略了证明。）在原点处由1和uv的张角等于1和u以及1和v张角之和（所有角度都按反时针方向测量），uv离开原点的距离是u和v离开原点距离的乘积。这可以等效地叙述为，由0，v和uv形成的三角形与由0，1和u形成的三角形相似，并且具有相同的指向。（精力充沛而对此不熟悉的读者也许可以利用早先给出的复数加法和乘法的代数规则以及上面的三角等式来直接证明这些结果。）

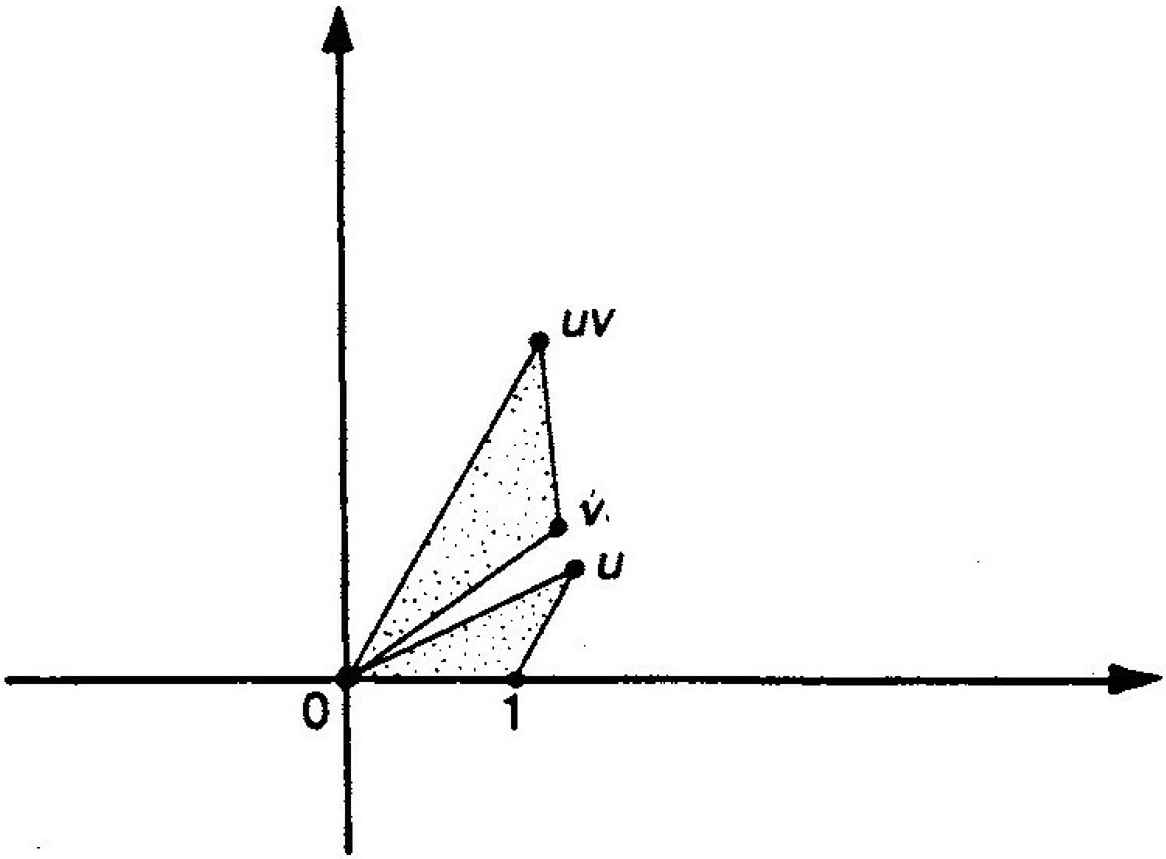


图3.11　两个复数u和v的乘积uv使得由0，u和uv形成的三角形与由0，1和u形成的相似，可以等效地说：uv到0的距离是u和v到0的距离的乘积，而uv和实轴（水平）构成的角度是u和v和该轴夹角的和

芒德布罗集的构成

我们现在可以看看如何定义芒德布罗集了。令z为一个任意选择的复数。不管这一个复数是什么，它都由亚根平面上的某一点所代表。现在考虑由下式

$$z \rightarrow z^2 + C$$

表出的映射，它把z由一个新的复数来取代。这儿C为另一个固定的（也就是给定的）复数。数z^2+C在亚根平面为某一个新的点所表示。例如，如果C刚好给出1.63-i4.2，则z就按点

$z \to z^2 + 1.63 - i4.2$

来映射。这样，特别是3就被

$3^2 + 1.63 - i4.2 = 9 + 1.63 - i4.2 = 10.63 - i4.2$

所取代，而-2.7+i0.3就会被

$(-2.7 + i\ 0.3)^2 + 1.63 - i\ 4.2 = (-2.7)^2 - (0.3)^2 + 1.63 + i\{2(-2.7)(0.3) - 4.2\} = 8.83 - i5.82$

所取代。当这些数变得复杂时，最好用电脑来进行这些计算。

现在不管C是多少，特别是点0在这个方案下被数C所取代。C本身又如何呢？它被C^2+C取代。假定我们继续这个步骤，将这种取代应用于C^2+C，则就得到

$(C^2+C)^2+C=C^4+2C^3+C^2+C.$

让我们再重复这个代换，把它应用到上面的数就得到

$$(C^4+2C^3+C^2+C)^2+C=$$

$$C^8+4C^7+6C^6+6C^5+5C^4+2C^3+C^2+C。$$

然后再对此数代换等。我们得到从0开始的一个序列

$$0，C，C^2+C，C^4+2C^3+C^2+C，\cdots。$$

现在如果我们选择一定的复数C来进行，则由这种办法得到的数的序列在亚根平面上永远不会徘徊到离原点非常远的地方去；更精确地

讲，对于C的这种选择该序列是有界的，也就是说序列的每一个成员都位于以原点为中心的某一个固定圆周之内（图3.12）。C=0的情况是一个好例子，由于在这种情形下序列的所有成员都是0。另一发生有界行为的例子是C=-1，因为此序列为0，-1，0，-1，0，-1，……。还有另一例子是C=i，其序列为0，i, i-1，-i, i-1，-i, i-1，-i，……。然而，对于其他不同的复数C，序列徘徊到离原点越来越远的不定距离的地方去；也就是说该序列是无界的，不能被包容于一个固定的圆周之内。这种行为的例子发生在当C=1时，因为这时序列变为0，1，2，5，26，677，458330，……。C=-3时也发生这种行为，其序列为0，-3，6，33，1086，……，还有C=i-1，序列为0，i-1，-i-1，-1+i3，-9-i5，55+i91，-5257+i1001，……。

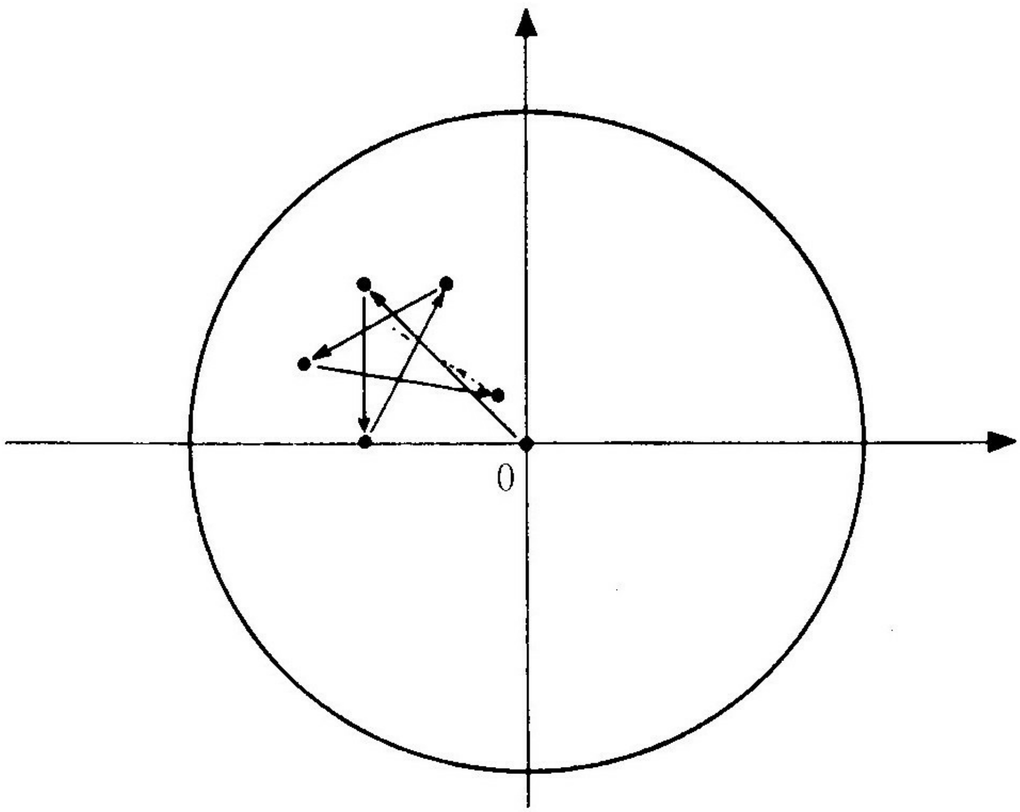


图3.12　如果在亚根平面上存在包括序列所有点的某一个固定圆周，则该序列是有界的（这个特殊的迭代从0开始并且C=−1/2+1/2i）

芒德布罗集，也就是我们托伯列南世界的黑色区域，正是亚根平面上由其序列维持有界的所有点C所组成的。白色区域是由产生无界序列的C所构成。我们前面所看到的细致的图像都是由电脑输出而绘成的。

电脑系统地跑过所有可能的复数C，并对任意选取的C算出序列0，C, C^2 +C，……，按照某种合适的判据来决定该序列发散与否。如果它是有界的，电脑就在屏幕上对应于C的那一点画上黑的。如果它是无界的，则画白的。电脑在所考虑区域的每一点都会最终决定画上白的或黑的颜色。

芒德布罗集的复杂性是非常引人注目的，尤其是和以下事实形成鲜明对照，这个集的定义在数学上是如此之简单。另外，这个集的一般结构对我们选取的$z \to z^2 + C$的映射的代数形式并不敏感。许多其他的递推的代数复映射（例如$z \to z^2 + iz^2 + C$）会给出极其类似的结构（假定我们从选取一个合适的数开始——也许不是零，对于每个适当选取的映射这一个数是按照一个明确的数学法则选取的）。就递推的复映射而言，这些“芒德布罗”结构的确有一种普适的绝对的特征。研究这种结构本身是数学中称作复动力系统的学科。

数学概念的柏拉图实在

数学家世界的对象有多“实在”？一种观点认为。它们似乎根本就没有任何是实在的。数学对象仅仅是概念；它们是数学家制造的精神上的理想化，它经常受到我们四周世界的外观和表面秩序的刺激，但充其量仍不过是精神的理想化而已。它们能不仅仅是人类头脑的恣意创造物吗？同时人们经常发现，这些数学概念会显示出某种深刻的实在性，完全超越出个别数学家的深思熟虑之外。人类思想恰如受到真理的引导，其真理本身具有实在性，而且只能对我们之中任何人揭示一部分真理。

芒德布罗集提供了一个突出的例子。它的美妙和复杂无比的结果既非任何人的发明，也不是任何一群数学家的设计。波兰——美国数学家（兼分形理论的领袖）贝内特·芒德布罗首先 [3] 研究了该集合。他对其中蕴含的美妙的细节并无预先的概念，尽管他知道正在寻找某种非常有趣的东西，的确，当他的第一张电脑画图开始出现时，他的印象是，所看到的模糊的结构只是电脑失误的结果（*Mandelbrot 1986* ）！他到了后来才相信集合就在那里。不但我们中的任何一个人都不能完全理解，而且任何电脑都不能指示芒德布罗集结构的复杂完整的细节。这个结构似乎不仅是我们精神的一部分，其本身也具有实在性。不管选择任一位数学家或任一台电脑去考察该集合，都会发现是对上述基本数学结构的

近似。用哪台电脑去进行计算都不会有真正的区别（假如电脑处于准确的工作状态），除了计算速度和存储与画图能力的差异会导致细节以及产生该细节的速度差别之外。使用电脑和在探索物理世界时使用实验仪器的方法在本质上是相同的。芒德布罗集不是人类思维的发明：它是一个发现。正如喜马拉雅山那样，芒德布罗集就在那里！

类似地，复数系统本身具有根本而永恒的实在性，它超越出任何特殊的数学家的精神构想。大致在杰罗拉莫·卡尔达诺的工作中复数才开始受到赏识。他是生于1501年死于1576年的意大利人，也是正式的医生、赌徒兼占星家（还为基督算过命）。1545年他写了一本重要的影响久远的代数专著《大术》。他在该书中首次提出了一般的立方方程的（以n次方根表达的）解的表达式[17]。然而，他注意到，在某一类方程具有3个实解的被人们称为“不可约”的情况下，在他的表达式的某一阶段必须取负数的平方根。虽然他为此深感迷惑，他却意识到，如果允许他取这种平方根，也只有这样，才能表达出全部答案（最后答案总是实的）。后来，1572年R.邦贝利在他题为《代数》的著作中，推广了卡尔达诺的结果并开始研究真正的复数代数。

初看起来，这样地引进负数的平方根似乎仅仅是作为工具——为了达到特定目的的数学发明——后来人们越来越清楚，从这些东西所获取的比原先所设计的多得多。正如我在前面提到的，虽然复数引进的当初目的是为了使取平方根畅通无阻，后来人们发现作为奖赏，能够求任何其他根式或者解任何代数方程。我们还发现了复数的许多神奇性质，这些最初一点儿的征兆也没有。这些性质现存在那里。尽管卡尔达诺、邦贝利、沃利斯、可提斯、欧拉、温塞尔和高斯具有无可怀疑的远见，这些性质不是由他们以及其他伟大的数学家放在那儿的。这些神奇是他们逐渐揭开的结构本身所固有的。当初卡尔达诺引进复数时，他根本对接踵而来的许多神奇没有任何一点暗示——而这些神奇的性质后来以不同的人来命名，例如柯西积分公式、黎曼映射定理以及卢伊扩张性质。这些以及其他显著的事实，正是卡尔达诺在1539年左右遭遇到的没有做过任何修正的那种数的性质。

数学究竟是发明还是发现？当数学家获得他们的结果时，是否仅仅产生了精神上的复杂构想没有客观实在性，但它们是这样的有力和精巧，甚至把发明者也愚弄了，并使他们相信这些仅仅为精神的构想是“实在的”？或者数学家实际上是发现现成的真理——这种真理的存在

完全独立于数学家的活动呢？我想到了现在，读者会很清楚，至少就复数的这种结构以及芒德布罗集而言，我执著地坚持第二种而不是第一种观点。

但是，情况也许还不像这么直截了当。正如我说过的，在数学中有些东西，用术语“发现”的确比“发明”更贴切得多，正如上面引用的例子。这些正是从结构出来的东西比预先放进的东西多得多的情形。人们可以认为，在这种情形下数学家和“上帝的杰作”邂逅。然而，还有其他情形，数学结构并没有如此令人信服的唯一性。例如，在证明某些结果的过程中，数学家发现必须引进某种巧妙的而同时远非唯一的构想，以得到某种特别的结果。在这种情形下，从构想得出的结果不太可能比起先放进的更多，所以术语“发明”似乎比“发现”更为妥当。这些的确只是“人的作品”。从这种观点看，真正的数学发现一般地被认为比“仅仅”发明具有更伟大的成就和抱负。

这种分类法在艺术和工程中是相当熟悉的。伟大的艺术作品的确比不甚伟大者“更接近于上帝”。在艺术家最伟大的作品中，揭示了某种预先的天界存在的[18] 不朽真理，而他们较差的作品可能更随意，但本质上只不过是会枯朽的作品，这种感觉对于艺术家并不稀罕。类似地，在漂亮组织的工程实施中，使用某些简单的预想不到的想法，并得到大量的成果，把这工程描述为发现比发明更妥当。

在叙述了这么多以后，我不禁感到，在数学中，至少对于其中某些最基本的概念，某种天国的不朽存在的信念比在其他情形下更强烈得多。在这种数学观念中存在比在艺术和工程中强烈得多的令人信服的唯一性和普适性。数学观念可在这样一种超越时间的天国的意义上存在的思想，是在古代（公元前360年左右）由伟大的希腊哲学家柏拉图提出的。随后这种思想就时常被称为数学柏拉图主义。它以后对我们很重要。

我在第1章用了一些篇幅讨论强人工智能的观点，根据这种观点，假设精神现象可在一个算法的数学观念中找到栖身之所。我在第2章中强调，算法的概念的确是根本的并为“上帝赋予”的思想。我在同一章论证道，这种“上帝赋予”的数学观念应有某种遗世独立的品格。由于为精神现象提供某种天界存在的可能性，该观点是否赋予强人工智能观点某些信任度呢？也许是这样的——我甚至在下面进一步作和这个观点相似

的推测。但是，如果精神现象的确可以找到这种一般的归宿，我不相信，这种归宿会是算法的概念。这里需要某种更微妙得多的东西。算法的东西只构成数学中非常狭小和有限的部分的这一事实是下面讨论的重要方面。我们将在下一章看到非算法数学的范畴和微妙之处。

第4章　真理、证明和洞察

数学的希尔伯特计划

什么是真理？我们如何对世界的真假形成判断呢？我们是否简单地遵循着某些算法？这种算法由于自然选择的强有力的过程无疑地比其他效率更低的可能算法更加优越。或许还有其他探索真理的非算法的途径——直觉、禀性或洞察。这似乎是一个困难的问题。我们的判断是基于感觉数据、推理和猜测的盘根错节的结合。而且，在世间的许多情势中也许并没有何为真何为假的共识。为了使问题简化，让我们只考虑数学真理。我们如何形成自己关于数学问题的判断或许“某些”知识呢？在这儿事情至少应该是更明了些。关于究竟什么为真什么为假在这里不应成为问题——难道会有问题吗？究竟什么是数学的真理呢？

数学的真理是一个非常古老的问题，这可回溯到早期的希腊哲学家和数学家的时代——并毫无疑问地比这还要更早。但是，只有在100多年前人们才刚刚获得了一些伟大的彻悟以及令人眼花缭乱的新的洞察。我们想要理解的正是这些非常基本的问题。它正好触及了我们的思维过程在本质上是不是完全算法的问题。议定这些问题是非常重要的。

数学在19世纪下半叶有了伟大的进展，其部分原因在于人们发展了数学证明的越来越有力的方法。（我们在前面提到的大卫·希尔伯特和乔治·康托尔，还有将要提到的伟大的法国数学家亨利·庞加莱是处于发展最前沿的3位。）数学家在利用如此有力的方法时相应地获得自信心。其中许多方法涉及去考虑具有无穷多元素的集合[19]。正是由于可能将这样的集合当成实在的“东西”——完全存在的整体，而不仅仅为潜在的存在，使证明经常得到成功。这许多强有力的观念是从康托尔的高度创造性的无穷数的概念中孕育而来的。他利用无穷集合系统地发展了这一切。（我们在上一章对此有所领略。）

然而，1902年英国逻辑学家兼哲学家贝特朗·罗素提出其著名的悖论，完全粉碎了这种自信心。（康托尔已预示过这一悖论，并且它是康托尔“对角线法”的直系后代。）为了理解罗素的论证，我们首先对把许

多集合当作完整的整体来考虑应有些了解。我们可以想象，某些集合是按照一个特殊的性质来表征的。例如，红的东西的集合是根据红性来表征的：就是说唯有当某物具有红性时才属于该集合。这样就允许我们把事情倒过来，按照单独对象也就是具有同一性质的事物的整个集合来谈论该性质。依照这种观点，“红性”是所有红的东西的集合。（我们还可以认为某一其他的集合就在“那里”，它们的元素为稍微复杂的性质所表征。）

这种按照集合定义概念的思想是1884年由具有影响的德国逻辑学家哥特洛伯·弗列格引进的步骤的核心。他可按照集合来定义数。例如，实际的数3是什么意思呢？我们知道“三性”是什么性质，但是3本身是什么？现在“三性”是一群对象的性质，也就是一个集合的性质：唯有如果当该集合不多不少有3个成员，则它具有“三性”的特别性质。例如，在特定的奥林匹克比赛中，奖章获得者的集合具有“三性”。还有三轮车的轮子集合，正常三叶草的叶的集合或者方程$x^3-6x^2+11x-6=0$的解的集合。那么，弗列格关于实在的数3的定义是什么呢？依照弗列格的论点，3必须是一个集合的集合：即所有具有“三性”[1]的集合的集合。这样，一个集合如果也只有如果属于弗列格集3，才具有3个成员。

这似乎显得有点啰唆，但实际上并非如此。我们可以把数一般地定义为对等集合的总体，这里对等的意思是讲“具有能一一配对的元素”（用通常的术语也就是“具有同样多的成员”）。数3就是这些集合的一个特例，其中的一个成员可以是包括一个苹果、一个橘子和一个梨的集合。请注意，这和丘奇在93页给出的“3”的定义完全不同。还可以给出其他今日相当流行的定义。

那么，罗素悖论又是怎么回事呢？它是关于以如下方式定义的集合R：

R为一切不是自身元素的集合的集合。

这样，R是集合的某一整体；集X属于该整体的判据是集X自身不是它自身的成员。

假定一个集合可以实际是它自身的一个成员，这是否非常荒谬？不见得。例如，考虑一个无穷集合（具有无穷多元素的集合）的集合I。

肯定存在无穷多不同的无穷集，这样I自身也是无穷的。这样I确实属于自身！那么，罗素的概念又如何导致悖论呢？我们问：罗素集合是它自身的一个成员或者不是它的成员？如果它不是它自身的成员，则它必须属于R，因为R刚好包括那些不是自身成员的集合。这样，R毕竟属于R—这是矛盾。另一方面，如果R是它的一个成员，那么由于“自身”实际上就是R，它就属于由自身并非其成员所表征的集合中，也就是它根本不是自身的成员——又导致矛盾[20]！

这种考虑并不轻率。罗素只不过以相当极端的形式利用数学家们正开始在证明中使用的、非常一般的、数学集论的同一类型的推理。事情很清楚地失去了控制，所以去弄清何种推理是允许的，何种是不允许的，应是适当的。很明显，可允许的推理必须没有冲突，而且只有真的陈述才能允许从原先已知的真的陈述中推导而来。罗素本人和他的合作者阿弗列德·诺斯·怀特黑德着手发展一种高度形式化的公理和步骤法则的数学系统，野心勃勃地要把所有正确的数学推理翻译到他们的规划中去。他们非常仔细地选择法则以防止导致罗素自己悖论的那种悖论的推理类型。罗素和怀特黑德所完成的业绩是一部纪念碑式的著作。然而，它是非常繁琐的，并且它实际上统一处理的数学推理的类型是相当有限的。我们在第2章首次提到的伟大的数学家大卫·希尔伯特致力于一个更可行更广泛的计划。它囊括了所有特殊领域的一切正确的数学推理类型。而且，希尔伯特倾向于认为，有可能证明该计划免于矛盾冲突。那么数学就一劳永逸地处于无可争辩的安全基础之上。

然而，1931年25岁的奥地利天才、数理逻辑学家库尔特·哥德尔提出了一个实质上摧毁了希尔伯特计划的令人震惊的定理，使得希尔伯特及其追随者的希望落空。哥德尔指出的是，任何精确（“形式的”）数学的公理和步骤法则系统，只要它大到足以包含简单算术命题的描述（诸如第2章考虑过的“费马大定理”），并且其中没有矛盾，则必然包含某些用该系统内所允许的方法既不能证实也不能证伪的陈述。这种陈述的真理性以可允许的步骤是“不能判定的”。事实上，哥德尔能够向我们证明，公理系统本身的协调性的陈述被编码成适当的算术命题后，必定成为一个这种“不能判定的”命题。理解这个“不可判定性”的性质对我们很重要。我们将要看到为何哥德尔的论证直接捣毁了希尔伯特计划的核心。我们还将看到哥德尔的论证如何使我们能用直觉去超越所考虑的任何个别的形式化的数学系统的局限。这一点理解对于下面大部分讨论至关重要。

形式数学系统

我们必须把“公理和步骤法则的形式数学系统”的含义弄得更清楚些。先必须假定有一符号表，我们的数学陈述用这些符号来表达。为了使算术能归并到该系统中去，这些符号必须足够于用来表示自然数。如果需要的话，我们可以只用通常的阿拉伯数的记号0，1，2，3，……9，10，11，12，……，虽然这使得法则的说明比所需要的稍微复杂一些。我们如果譬如讲用0，01，011，0111，01111，……去表示自然数列（或作为折中，我们可以用二进制记数法），则说明就会简单得多。然而，由于这会在以下的讨论中引起混淆，所以在我的描述中只用通常的阿拉伯记号，而不管系统在实际上用什么符号。我们也许需要一个“间隔”符号去把我们系统的不同的“词”或“数”分开，但这又是令人混淆的，所以为了必要的目的我们可以只用（，）。我们还需要用字母来表示任意（“变量”）自然数（或许整数、分数等——但是让我们在这里只局限于自然数），譬如t, u，v, w，x, y，z, t'，t''，t'''，……。符号t'，t''……也许是需要的，因为我们不想对表式中可能出现的变量数目加上一个上限。我们把（'）当做形式系统的另外的符号，这样使符号实际数目保持为有限。我们还需要基本算术运算的符号=，+，×，等等，也许还需要不同种类的括号（，），[，]以及诸如&（“和”），⇒（“蕴涵”），∨（“或”），⇔（“当且仅当”），～（“非”，或“不是以下论述”）等逻辑符号。此外我们还需要逻辑的“量词”：存在量词∃（“存在……使得”）和全称量词[∀]（“对于每一个……我们有”）。现在，我们可以把诸如“费马大定理”的陈述写成：

$$\sim\exists w, x, y, z[(x+1)^{w+3}+(y+1)^{w+3}=(z+1)^{w+3}]$$

（见第2章76页）。（我原可以用0111来表示3，或者利用“自乘”的记号使得和形式化符合得更好；但是正如我说过的，我只拘泥于传统的符号，以避免引起不必要的混淆。）上面的陈述（到第一方括号处结果）的意思为：

“不存在自然数w, x，y, z使得……”。

我们还可以用∀把“费马大定理”重写成：

$$\forall w, x, y, z[\sim(x+1)^{w+3}+(y+1)^{w+3}=(z+1)^{w+3}],$$

其意思（到第一方括号的“非”符号处结束）为：

“对于所有的自然数w, x，y, z下述不真……”。

这和前面在逻辑上是相同的。

我们需用字母来表示整个命题，为此目的我用大写字母P,　Q，R, S，……。如下的一个命题事实上为上面的费马的论断：

$F=\sim\exists w, x, y, z[(x+1)^{w+3}+(y+1)^{w+3}=(z+1)^{w+3}]$。

一个命题也可依赖于一个或更多的变量；例如，我们也许对某一特殊的[21]幂指数w+3下的费马论断感兴趣：

$G(w)=\sim\exists x, y, z[(x+1)^{w+3}+(y+1)^{w+3}=(z+1)^{w+3}]$,

这样G（0）断言“没有一个立方可代表正数立方之和”，G（1）对四次方作同样断定，等等。（注意∃之后的w没有出现。）现在费马论断是说，G（w）对所有的w成立。

$F=[\forall]w[G(w)]$。

G（）是一个所谓的命题函数，也就是依赖于一个或多个变量的命题的例子。

系统的公理是由一般命题的有限罗列所构成，假定在符号的意义已给定的情形下，这些命题的真理性是不证自明的。例如，对于任何命题或命题函数P, Q，R（），在我们公理之中有

$(P\&Q)\Rightarrow P$,

$\sim(\sim P)\Leftrightarrow P$,

$\sim\exists x[R(x)]\Leftrightarrow\forall x[\sim R(x)]$,

其“自明的真理性”清楚地可由其意义所确定。（第一个简单地断定：“如果P和Q都为真，那么P为真”；第二个断定：“P不真的断言为不真”和“P为真”是等价的；第三个可用上面给出的“费马大定理”的两种叙

述方法的逻辑等价性作为例子）。我们还可包括基本的算术公理，诸如

$\forall x, y[x+y=y+x]$

$\forall x, y，z[（x+y）\times z=（x\times z）+（y\times z）]$，

尽管人们也许宁愿从某些更初等的东西建立这些算术运算，并将这些陈述作为定理导出。步骤法则是诸如这样（自明）的东西：

“从P和P⇒Q我们可推出Q”，

“从∀x[R（x）]我们可推出把一特殊的自然数代入到R（x）中的x而得出的任何命题”。

这些是告诉我们如何从已成立的命题引出新命题的方针。

现在从公理开始，然后不断重复应用步骤法则，就可以建立起一长串的命题。我们在任何阶段都可再使用这些公理，并且总可以不断使用任何我们已经添加到越来越长的表上的命题。任何正确地集合到表上的命题都被称作定理（虽然它们中有许多是相当无聊和无趣的）。如果我们有一个要证明的特定的命题P，则我们可去找一个表，这个表按照这些法则正确地集合起来，并用我们特定的命题P作为终结。这样的表在我们的系统中为我们提供了一个P的证明；而P就相应地成为一个定理。

希尔伯特计划的思想是，对于任何定义好的数学领域，去找一足够广泛的公理和步骤法则的表，使得所有适合于该领域的正确的数学推理的形式都可以编入。让我们把数学领域暂定为算术（包括量词∃和∀使得可以作诸如“费马大定理”的陈述）。考虑比这更一般的数学领域在这里对我们并无益处。算术已经是足够一般到可以应用哥德尔步骤的地步。如果我们能够接受这样的事实，即这个公理和步骤法则的广泛系统，按照希尔伯特计划，的确把算术给予我们，那么它就为我们提供对算术中任何命题数学证明的“正确性”的确定判据。人们存在过希望，这样的公理和法则系统也许是完备的，也就是它会使我们在原则上决定任何可在此系统中表述的数学陈述的真伪。

希尔伯特的希望是，对于任何一串代表一个数学命题的符号，譬如讲P，人们应能证明P或者～P，依P是真的还是假的而定。我们在这里

必须假定该符号串在构造上是语法正确的，也就是满足所有形式主义的记号法则，诸如括号必须正确地配对等——使得P具有定义清楚的真的或假的意义。如果希尔伯特的希望能被实现，这甚至使我们不必为这些命题的意义忧虑！P仅仅为一语法正确的符号串。如果P为一道定理（也就是可在系统内证明P），则符号串P的真值就可被赋予真。另一方面，如果能证明～P为定理的话，则可被赋予假。为了使这些有意义，我们除了完备性外还需要一致性。也就是说，不应有P和～P都为定理的符号串P。否则P会同时是真的和假的！

把数学陈述中的意义抽走，只把它们当成某种形式数学系统的符号串是形式主义的数学观点。有些人喜欢这种观点，而数学就变成一种“无意义的游戏”。然而，我不欣赏这种观点。确实是“意义”而非盲目的算法计算才赋予数学以实质。庆幸的是，哥德尔给了形式主义以毁灭性的打击！让我们看看他是怎么做的！

哥德尔定理

哥德尔论证的部分是非常繁琐和复杂的。然而我们没有必要去考察那纷乱的部分。另一方面，其中心思想是简单、漂亮和深刻的。这就是我们可能鉴赏的部分。其复杂的部分（其中不乏许多巧妙之处）仔细说明如何把形式系统的个别步骤法则以及不同公理的使用实际地编码成算术运算。（意识到这是一个富有成果的可进行的工作正是其深刻部分的一个方面！）为了实现编码，人们需要找到用自然数来对命题编号的某种方便方式。一种方法就是简单地对形式系统每个特定长度的符号串使用某种“字典”顺序，按照串的长度还有一个总的顺序。（这样，长度为1的串可按字母顺序排列，接着的是按字母顺序排列的长度为2的串，再后面是长度为3的串等）。这叫作字典顺序[22]。哥德尔原先用的编号顺序更复杂，但是这种差异对我们不重要。我们将特别关心依赖于单变量的命题函数，譬如上述的G（w）。令应用于w的第n个这样的命题函数（在选定的符号串顺序下）为：

P_n（w）.

如果我们愿意的话，可以让编号稍微有点“草率”，这样我们的一些表式可能语法上不正确。（这可使算术编码比在试图略去这种语法不正

确的表式时容易得多。）如果P_n（w）是语法正确的，它就是关于两个自然数n和w的定义好的特定的算术陈述。哪一个算术陈述是

准确的应依所选取的特定编号系统的细节而定。那是属于论证的复杂部分，在此不予关心。构成系统中的某一定理的证明的一串命题在选定的编序方案中也可用自然数编号。令

$\prod_n$

表示第n个证明。（这里我又一次使用“草率的编号”，对于某些n的值，可能表示式“$\prod_n$”的语法不正确，并因此没有证明什么定理。）

现在考虑如下的依赖于自然数w的命题函数

～∃x[$\prod_n$证明P_w（w）]。

在方括号中的陈述的一部分使用了文字，但它是完全精确定义的。它断定第x个证明实际上是P_w（）应用于值w本身的命题的证明。方括号之外的被否定的存在量衡用以移走一个变量（“不存在一个x使得……”），这样我们得到了一个只依赖于一个变量w的算术的命题函数。此整个表达式断定不存在P_w（w）的证明。我假定它的语法是正确的[甚至如果P_w（w）的语法不正确——在这种情形下该陈述仍然是对的，因为一个语法错误的表达式是不能被证明的]。由于事实上我们已假设将其转换成算术，所以上面实际上是关于自然数的某一算术的陈述（方括号中的部分为定义得很好的关于两个自然数x和w的算术描述）。该陈述是可以被编码成算术，但这一点并不假设是明显的。说明这样的陈述的确可被编码，涉及哥德尔论证的复杂部分的主要“困难工作”。正和前面一样，它究竟为哪个算术陈述将依赖于编号系统的细节，并大大地依赖于我们形式系统的公理和法则的结构细节。由于所有那些都属于复杂的部分，我们在这里不关心其细节。

我们已将所有依赖于单变量的命题函数编号，所以我们刚刚写下的必须赋予一个数。让我们把这个数记作k。我们的命题函数是在表上的第k个。这样：

～∃x[∏x证明P_w（w）]=P_k（w）。

现在对特殊的w值即w=k来考察这一个函数。我们得到：

$\sim\exists x[\prod_x$ 证明P_k （k）$]=P_k$ （k）。

这个特定的命题P_k （k）是完好定义（语法正确）的算术陈述。它是否可在我们形式系统中有一个证明呢？它的反命题$\sim P_k$ （k）有证明吗？这两个问题的答案都是“否”。从考察作为哥德尔步骤基础的意义可以看到这一点。虽然P_k （k）仅仅是一个算术命题，但我们已经将其构造成，使得写在左边的论断为“在这系统中不存在命题P_k （k）的证明”。如果我们非常仔细地设定好我们的公理和步骤法则，并假定做了正确的编号，则在这系统中不能存在这个P_k （k）的证明。因为如果存在这样的证明，则P_k （k）实际断言的陈述的意义，也就是不存在证明，将是错的，这样作为一个算术命题的P_k （k）就必须是错的。我们的形式系统不应构造得这么坏，使得它在实际上去允许证明错的命题！所以情况只能是P_k （k）在事实上无法证明。而这正是P_k （k）要告诉我们的。所以断定P_k （k）必须是一真的陈述，这样P_k （k）作为算术命题必须为真。于是，我们已经发现了在该系统中没有证明的真的命题！

关于它的反命题$\sim P_k$ （k）我们可以说些什么呢？最好我们也不能找到它的证明。我们刚刚建立了$\sim P_k$ （k）必须是错的[因为P_k （k）是真的]，而我们假定不能在此系统中证明错的命题！这样无论P_k （k）还是$\sim P_k$ （k）在我们的形式系统中都是不可证明的。哥德尔定理就这样地被建立起来了。

数学洞察

请注意，在这里发生了某种非常奇异的事情。人们经常把哥德尔定理当作某种负面的东西——显示了形式化数学推理的不可避免的局限性。不管我们自以为是多么有智慧，总有些命题漏网。但是，我们是否要为这一特殊的命题P_k （k）忧虑呢？在上述的论证过程中，事实上我们已建立了P_k （k）是一个真的陈述！尽管在该系统中不能形式地证明

这个事实，不管怎么样我们已设法看到了这一点。真正需要忧虑的人倒是严格的数学形式主义者。这是因为从这推理我们已确定形式主义者的“真理”概念不可避免地是不完备的。不管把哪一个（一致的）形式系统应用于算术，总存在一些命题我们可以看到是真的，但用形式主义者提出的上述过程不能赋予真理值为真的命题。一个严格的形式主义者试图躲开这个情况的可能方法也许是根本不提真理的概念，而仅仅讲在某一固定的形式系统中的可证明性。然而，这显得非常局限。由于哥德尔论证的基本点利用关于何者实际上为真的何者不真的推理，人们甚至都不能作出上述的论证[2]。一些形式主义者采用更“程序化”的观点，断言不去忧虑诸如P_k（k）这样的陈述，由于它们作为算术命题来讲极端复杂和乏味。这些人会宣称：

是的，存在一些诸如P_k（k）的古怪的陈述，对于这些陈述我的可证明性或真理的概念不和你们的真理的内禀概念相符合。但是那些陈述却不会在严肃的（至少在我所感兴趣的那种）数学中出现。这是因为作为数学而言，这样的陈述是荒谬绝伦地复杂和不自然。

的确，像P_k（k）这样的作为关于数的数学描述的命题，被全部写出时，会是极端繁琐和古怪的。但是近年来，人们提出了一些具有非常可接受特性的相当简单的陈述，它们实际上等价于哥德尔类型的命题[3]。这些命题不能从正常的算术公理得到证明，而是从公理系统本身所具有的“显然正确”的性质而来。

对我来讲，形式主义者对“数学真理”缺乏职业的兴趣，似乎是对数学哲学所采取的非常古怪的观点。而且，也确实不是那么切合实际。当数学家在进行推理时，他们没必要继续不断地检查他们的论证是否可按照某个复杂的形式系统的公理和步骤法则来表达。他们只要肯定其论证是确定真理的有效方法即可。哥德尔的论证是另一类有效步骤。这样我似乎认为，P_k（k）正和能利用预先给出的公理和步骤法则更传统地得到的数学真理一样好。

建议进行如下步骤。我们把P_k（k）接受为真正有效的命题，并简单地表示为G_0；这样可以把它作为一个额外的公理加到系统中去。当然，我们新的修改的系统又有了它自己的哥德尔命题，譬如讲G_1，它又是一个完全有效的关于数的描述。我们相应地又把G_1加到我们的系

统，由此得到进一步修改的系统，它又有自己的哥德尔命题G_2 （又是完全有效的），我们又把它合并进去，得到了下一个哥德尔命题G_3 ，再合并等，无限次地重复这一过程。当我们允许使用整列的G_0 ，G_1 ，G_2 ，G_3 ，……作为附加的公理时，结果的系统是什么呢？它可以是完备的吗？由于现在我们有了一个无限制（无限）的公理系统，哥德尔步骤能否适用也许不太清楚。然而，不断附加哥德尔命题是一个完全系统化的方案，我们可将其当作通常的公理和步骤法则的有限的逻辑系统来重述。这一系统又有它自己的哥德尔命题，譬如讲G_w ，它又能被用来作为公理去附加，而形成了所得到的系统的哥德尔命题G_{w+1} 。正如上面那样重复，我们得到了命题G_w ，G_{w+1} ，G_{w+2} ，G_{w+3} ，……的表，所有都是关于自然数的完全有效的陈述，并可附加到我们的形式系统中去。这又是完全系统化的，它导致一个包罗这一切的新系统；但是它又有自己的哥德尔命题，譬如讲G_{w+w} ，我们可将其重写成G_{w2} 。而整个步骤又可重新开始，我们得到一个新的无穷的，却是系统的公理G_{w2} ，$G\ w_{2+1}$ ，G_{w2+2} 等的表，它又导致一个新的系统以及一个新的哥德尔命题G_{w3} 。重复这整个过程，我们得到G_{w4} ，然后还有G_{w5} 等。现在这一步骤又是完全系统化的，并具有自身的哥德尔命题G_{w2} 。

这会有终结吗？在一种意义上讲没有；但它导致我们进入不能在此作细致讨论的某些困难的数学考虑。1939年阿伦·图灵在一篇论文[4] 中讨论了上面的步骤。事实上，令人印象深刻的是，任何真的（但普适量化的）算术命题都可由这类重复的“哥德尔化”步骤得到！可参阅[Feferman 1988] 。然而，这在一定程度上依赖于我们如何实际上决定一个命题真假的问题。在每一阶段关键的问题是如何把哥德尔命题的无穷族合并，从而提供一个单独的（或有限数目的）附加公理。这就要求我们的无穷族能以某种算术的方式被系统化。为了保证正确地完成所预想的系统化，我们要使用系统之外的直觉——正如我们首先为了看到P_k （k）是一个真的命题所做的那样。正是这些直觉是不能被系统化的——它必须超越于任何算法行为！

我们利用直觉得出哥德尔命题P_k （k）实际上是算术中的真的陈述，是被逻辑学家称之为反思原理步骤的普遍类型的一个例子：这样，由“反思”公理系统和步骤法则的意义，并使自己坚信这些的确是得到数

学真理的有效方法，人们可能把这直觉编码成进一步的、真的、不能从那些公理和法则推导出来的数学陈述。正如上面概述的，推出P_k （k）的真理性依赖于这样的一个原则。另一个与原先哥德尔论证相关（虽然在上面没提及）的反思原则依赖于如下的事实去推出新的数学真理，即我们已经相信能有效得到数学真理的公理系统实际上是协调的。反思原理经常涉及有关无穷集合的推理，人们使用的时候一定要小心，不要过于接近会导致罗素类型悖论的论证。反思原理为形式主义推理提供了反题。如果人们很小心的话，就能使他跳出任何形式系统的严格限制之外，并得到原先似乎得不到的新的数学洞察。在我们的数学文献中会有许多完全可接受的结果，其证明需要远远超越原先的算术标准形式系统的法则和公理的洞察。所有这些表明，数学家得到真理判断的心理过程，不能简单地归结为某个特别形式系统的步骤。虽然我们不能从公理推出哥德尔命题P_k （k），却能看到其有效性。这类涉及反思原理的“看见”需要数学的洞察力，而洞察不是能编码成某种数学形式系统的纯粹算法运算的结果。我们将在第10章再回到这个论题上来。

读者也许会注意到在建立P_k （k）“不可证明性”的真理和罗素悖论的论证之间的相似性，还和图灵解决停机问题的图灵机不存在的论证也有相似性。这些相似性不是偶然的。在这三者之间存在有强大的历史连接的脉络。图灵是在研习哥德尔工作之后才找到它的论证的。哥德尔本人非常熟悉罗素悖论，并能把这一类将逻辑延伸得这么远的悖论的推理转化成有效的数学论证。（所有这一切论证都起源于前一章描述的康托尔的“对角线法”。）

为什么我们应该接受哥德尔和图灵的论证，而必须排斥导致罗素悖论的推理呢？前者更直接明了得多，作为数学论证而言更出人意料，而罗素悖论则依靠牵涉到“巨大”集合的更为模糊的推理。但是必须承认，其差别并不真像人们以为的那么清楚。弄清这些差别的企图是整个形式主义观念的强大动机。哥德尔的论断表明，严格的形式主义者的观点是不能成立的，但他没有向我们指出另外完整的可信赖的观点。我认为这问题仍未解决。当代数学中为了避免导致罗素悖论的“巨大的”集合的推理的类型所实际采用 [23] 的步骤是不能完全令人满意的。而且，它仍然试图以明晰的形式主义的术语来表达，换句话说，按照我们并不完全相信不会出现矛盾的术语来描述。

无论情况如何，依我看来，哥德尔论证的清楚推论是，数学真理的概念不能包容于任何形式主义的框架之中。数学真理是某种超越纯粹形式主义的东西。甚至即使没有哥德尔定理，这一点也是清楚的。在我们去建立一个形式系统任何试图中，如何决定采取什么公理和步骤法则呢？我们在决定采取法则的指导总是，在给定系统的符号的“意义”下对何为“自明正确”的直觉理解。根据关于“自明”和“意义”的直观理解，我们如何决定采用哪个形式系统是有意义的，哪个是没意义的呢？以自身具有一贯性的概念来决定当然不够。人们可以有许多自身具有一贯性但在含义上没有“意义”的系统，它们的公理和步骤法则具有错误的意义，或者根本没有意义。甚至在没有哥德尔定理时，“自明”和“意义”的概念仍然是需要的。

然而，若没有哥德尔定理，人们可能想象“自明”和“意义”的直觉概念只要在开始建立形式系统时用一次就好了，而此后就与决定真理的清楚的数学论证不相干。那么按照形式主义者的观点，这些“模糊的”直觉概念在发现适当形式的论证时，作为数学的初步思维或者导引而起作用，而在实际展示数学真理时不起作用。哥德尔定理表明，这个观点在数学基本哲学中不能真正站住脚。数学真理的观念远远超越形式主义的整个概念。关于数学真理存在某些绝对的“上帝赋予”的东西。这就是在上一章结尾处讨论的柏拉图主义。任何特定的形式系统都具有临时和“人为”的品格，在数学的讨论中，这类系统的确起着非常有价值的作用，但是它只能为真理提供部分（或近似）的导引。真正的数学真理超越于人为的构造之外。

柏拉图主义或直觉主义

我已指出了数学哲学的两个相反的学派，我强烈地赞成柏拉图主义，而不赞成形式主义观点。我的划分实际上是非常朴素的。可以对此观点进行许多细致的推敲。例如，人们可以争论在“柏拉图主义”的总名称下，数学思维的对象是否为任何实际的“存在”，或者它只是绝对的数学“真理”的概念，我不想在此做任何鉴别。依我看来，数学真理的绝对性和数学概念的柏拉图存在性本质上是等同的一件事。例如，必须归于芒德布罗集的“存在”是其“绝对”性质的特征。亚根平面上的一点是否属于芒德布罗集是一个绝对的问题，与哪个数学家哪台电脑在作考察无关。正是芒德布罗集的“数学家无关性”赋予它柏拉图式的存在。而且，

它最精细的细节超过了我们目前使用电脑所能得到的极限。那些仪器只能得到具有更深刻的自身的“电脑无关”存在结构的近似。然而，我很欣赏对此问题的许多其他合情理的观点。在此我们不必过于忧虑这些差别。

如果的确有人声称自己为柏拉图主义者，他究竟愿意把柏拉图主义贯彻到何等程度，也有观点上的不同。哥德尔本人是一个非常强烈的柏拉图主义者。我迄今所考虑的数学陈述的类型是相当“缓和的”[5]。特别在集论中可引入更令人争议的陈述。当考虑集论的所有分支时，就会遭遇到构造极其庞大的模糊的集合，以至于像我这样坚定的柏拉图主义者都会怀疑其存在或它为“绝对的”东西[6]。也许会面临着这样的阶段，集合具有如此繁复以及概念上可疑的定义，以至于有关它们数学陈述的真假问题开始具有某种“个人品位”而非“上帝赋予”的品质。人们是否准备和哥德尔一道把柏拉图主义坚持到底，要求关于这么巨大集合的数学论述的真假总为一个绝对的或“柏拉图”的事体，或者人们在某处停止，只有当集合为合理地构成并且没有这么巨大时才寻求绝对的真假的解答，对我们的讨论关系并不重大。以我刚刚提到的标准看，对于我们具有意义的（有限或无限）集合，真是不可思议的微小！这样我们不必关心在这些不同柏拉图主义观点之间的差异。

然而，存在诸如称为直觉主义（或称作有限主义）的其他数学观点，它走到拒绝任何无穷集合的完整存在的另一极端[24]。直觉主义是1924年由荷兰数学家L.E.J.布劳威尔作为对某些（诸如罗素的）悖论的与形式主义相区别的响应而倡导的。这些悖论是由于在数学推理中太过自由地应用无穷集合所引起的。这种观点的根源可追溯到亚里士多德。他虽然是柏拉图的学生，却否定柏拉图关于数学本体的绝对存在和无穷集合的可接受性。直觉主义否认（无穷或其他）集合自身的“存在性”，而集合仅仅被当作可能确定其成员的规则。

布劳威尔的直觉主义的一个特征是排斥“排中律”。该定律宣称，一个陈述的否定之否定等效于该陈述。（可用符号表示为～（～P）⇔P，这是我们上面遇到的关系。）也许亚里士多德会对在逻辑上如此“显明的”东西受到排斥感到不悦！排中律按照“常识”被认为是自明的真理：如果某事物不真的断言是错的，则该事物一定是真的！（这一个定律是被称作反证法的数学方法的基础，参阅78页。）但是直觉主义者发现他们能推翻这一个定律。这基本上是因为他们对存在的概念采取不同的看

法，他们要求一个确定的（智力上的）构造必须是数学对象实际存在性被接受的先决条件。这样，对于直观主义者来说，“存在”的意思是“构造存在”。在一个用反证法来进行的数学论证中，人们提出某种假设，试图去显示出它的推论会导致一个矛盾，这个矛盾为问题中假设的谬误提供了所需的证明。此假设可采用这样的一个陈述，具有某些必需性质的数学事体不存在。当这个陈述导致矛盾时，在通常数学中，他就推论说所需的事体的确存在。但是，这样的论证本身并没为实际构造这样的事体提供任何手段。对于直觉主义者来说，这类存在根本就不是存在。他们正是在这个意义上拒绝接受排中律以及反证法的步骤。布劳威尔对此非构造性的“存在”深为不满[7]。他断言，没有一个实在的构造，这种存在的概念是无意义的。在布劳威尔的逻辑中，人们不能从某种对象的不存在性的谬误推导出该物体实际上的存在！

我认为，虽然关于从数学的存在中寻求构造有某些令人赞赏的东西，但布劳威尔的观点是过于极端了。布劳威尔在1924年首次提出他的思想，比丘奇和图灵的工作早十多年。现在按照图灵的可计算性的构造性概念可在数学哲学的传统框架内研究，并没有必要走到像布劳威尔那么极端的程度。我们可以把构造性的问题和数学存在性的问题分开来讨论。如果我们跟随直觉主义，就必须摒弃数学中非常强有力的论证的使用，而课题就变得有点窒息和虚弱。

我不想细述直觉主义观点导致的种种困难的荒谬；但是仅仅提及一些问题也许是有益的。布劳威尔经常关心提及的一个例子是π的小数展开：

3.141592653589793　…….

是否在这个展开的某一处存在20个接连的7的序列，也就是

π=3.141592653589793……77777777777777777777……

或者不存在这种情形呢？按照通常的数学，现在所有能说的是，或者存在或者不存在——而我们不知哪个是对的！这看来是一个肯定无害的描述。然而，除非人们已经（以某种直觉主义者接受的构造方式）确立存在这个序列或者不存在这个序列，他们实际上对讲“或者π的小数展开中某处存在连续20个7的序列或者不存在”采取否决的态度！直接的计

算也许足以显示在π的小数展开的某处的确存在20个连续的7的序列，但要确证没有这样的序列则需要某种数学定理。迄今电脑在计算π时还不能进行足够远到能确认该序列的存在。在基于概率的基础上，人们预料这样的序列的确存在。但是即使利用一台每秒能恒定产生10^{10} 位数的电脑，大约也需要100年或1000年左右才能找到这序列！我认为更可能是，不进行直接计算，该序列的存在某天会在数学上被确认（也许是作为推论某种更有力和更有趣得多的结果）——虽然也许不是以直觉主义者能接受的方式！

这一个特殊问题并不具有实际的数学趣味。它只是由于容易叙述才作为例子提出。以布劳威尔的直觉主义的极端形式，他会宣称：现在断言“在π的小数展开中的某处存在20位连续的7的序列”既不是真的亦不是假的。如果在将来用计算或（直觉主义的）数学证明得到适当的这种或那种结果，那么断言就变成“真”的或“假”的，视当时情况而定。“费马大定理”是一类似的例子。根据布劳威尔的极端直觉主义，现在这一个命题既不是真的亦不是假的，但将来也许会变成其中的一种。对我来讲，数学真理的这种主观性和时间依赖性是不可理喻的。数学结果是否或何时被接受为正式“证明了”的的确是一个主观的事体。但是数学真理不应取决于这些依赖社会的判据。对于人们希望能可靠地用来描述物理世界的数学，具有随时间而变的真理概念至少可以说是尴尬的和不令人满意的。并非所有的直觉主义者都采用布劳威尔那样强烈的观点。尽管这样，甚至对于那些同情构造主义目的的人也是这么认为，直觉主义观点显然是粗劣的。就仅仅因为人们可允许使用的数学推理的类型过于局限的原因，很少当代数学家愿意全心全意地追随直觉主义。

我已经简介了当代数学哲学的3个主流：形式主义、柏拉图主义和直觉主义。我并不掩饰自己强烈同情柏拉图主义的观点，也就是数学真理是绝对的、外在的、永恒的，并不基于人造的判据之上；数学对象具有超越时间的自身的存在，既不依赖于人类社会，也不依赖于特定的物体。我把这种观点贯穿于本节、上一节以及第3章的结尾处。我希望读者准备在这一点上和我大致“同心同德”。它对于后面要遇到的大量内容都很重要。

从图灵结果到类哥德尔定理

我在阐明哥德尔定理时忽略了许多细节，并且也忽略它的论证中或许在历史上的最重要的部分；这就是被叫作公理相容性的“不可判定性”。我在这里的目的不在于强调这“公理相容性的可证明性的问题”。这个问题对于希尔伯特及其同代人是如此之重要。我只是表明，利用所考虑的形式系统的公理和法则，某个特殊的哥德尔命题既不是可证明的也不是可证伪的。但是利用我们对该问题中运算意义的直觉可以清楚地看到，它是一个真的命题！

我提到过，图灵在研究了哥德尔的著作后发展了自己后来的论证，以确立停机问题的不可解性。这两个论证有许多共同的地方，事实上，哥德尔结果的关键方面可利用图灵步骤直接推出。让我们看看这是如何进行的，并因此对哥德尔定理的背后的东西有某种不同的洞察。

一个形式数学系统的主要性质是，决定某一给定的符号串是否构成该系统中给定的数学论断的证明应是可计算的事体。表达数学证明的全部要点毕竟在于对于什么是有效推理、什么是无效推理不必做进一步的裁决。以完全机械的和原先预定的办法来检查一个想象的证明是否确实是一个证明应是可能的；也就是说必须有检查证明的算法。另一方面，为提出的数学陈述去找证明（或证伪），我们并不要求它必须是算法的事。

事实上，在任何形式系统中只要某种证明存在，就总有找到证明的算法。由于我们必须假定该系统是以某种符号语言来表达的，这种语言是按照符号的某些有限“字母”来表达的。正如以前一样，让我们把符号串以字典的方式编序。我们记得这表示对于固定的串的长度按字母编序，先取所有串长为1的，然后串长为2的，串长为3的等（见139页）。这样，我们就把所有正确建立起来的证明按照这个字典方案进行编序。我们有了证明的列表，也就有了该形式系统的所有定理的列表。这是因为定理刚好是出现在正确构造的证明的最后一行的命题。这种列表完全是可计算的：由于不管系统的符号串是否有作为证明的意义，可以先考虑所有的串的字典列表，然后用我们的证明检查算法去检验其是否为一个证明，若不是则抛弃之；然后以同一方法检验第二个，若不是证明则抛弃之；然后第三、第四等。如果有一个证明，我们则可用这种办法最终在这一列表的某一处找到它。

这样，如果希尔伯特已经成功地找到它的公理和步骤法则的数学系

统。该系统足够有力到能使人们用形式证明决定任何在该系统中正确表达的数学命题的真假——则就会有一般的算法方法去决定任何这种命题的真理性。为什么会这样呢？因为用上述的步骤，如果最终在某个证明的最后一行遇到了我们所寻求的命题，则我们就证明了该命题。反之，如果我们最终遇到的一行是我们命题的否定，则我们就证伪了它。如果希尔伯特计划是完备的，这种或那种的终局就总会发生（并且，如果是协调的，两者永远不会同时发生）。这样，我们的机械步骤总会在某一阶段结束，而我们就应有一种决定系统所有命题真假的普通算法。这就和第2章阐述的图灵结果相冲突，也就是说不存在决定数学命题的一般算法。因而我们实际上证明了哥德尔定理，就是说希尔伯特期望的计划在刚刚讨论的意义上不可能是完备的。

由于哥德尔所关心的形式系统的类型只对算术命题而不是对一般的数学命题适用，所以事实上哥德尔定理比上述的更特定。我们是否能安排只用算术的运算去实现图灵机的所有必需的运算呢？换句话说，是否所有自然数的可计算功能（也就是图灵机动作的结果，递归的或算法的功能）可按通常的算术表达呢？我们几乎真的可以，但还不是。我们需要在算术和逻辑（包括∃和∀）的标准法则外加上一个额外的运算。这个运算简单地选择为：

“使得K（x）成立的最小自然数x”，

这儿K（）是任何给出的算术地可计算的命题函数——并假定存在这样的一个数，也就是∃x[K（x）]为真的。（如果没有这样的一个数，则我们的运算在试图寻求所需的不存在的x时就会“无限地算下去”[25]。）无论如何，在图灵结果的基础上前面的论证确认了，把数学的一切分支归结为某个形式系统中的计算的希尔伯特计划的确是不成立的。

就此而言，这一步骤并没有这么清楚地显示，在这系统中我们具有真的，但不能证明的一个哥德尔命题[例如P_k（k）]。然而，如果我们回忆在第2章给出的关于“如何超越算法”（参阅85页）的论证，我们就看到了可以做非常类似的事情。我们在那个论证中指出，如果有决定图灵机动作是否停止的任何算法，我们便能制造图灵机的一个动作，我们看到该动作不停止，但是该算法看不到这一点。（记得我们强调过，当一台图灵机将要停止时，该算法必须正确地通知我们，虽然在图灵机动

作不停止——它会永远运行下去的情形，有时它不能告诉我们。）鉴于上述的哥德尔定理的情形，我们具有利用洞察可以看到实际上必须为真的命题（图灵机的不停运行），但是给定的算法动作不能告诉我们这些。

递归可列集

存在一种按照集论的语言形象地描述图灵和哥德尔基本结果的方法。这就使得我们可以不用按照特别的符号主义或形式系统的任意描述，而使本质问题呈现出来。我们将只考虑（有穷或无穷的）自然数的集合0，1，2，3，4，……。这样我们将考察这些集合，诸如{4，5，8}，{0，57，100003}，{6}，{0}，{1，2，3，4，……，9999}，{1，2，3，4，……}，{0，2，4，6，8，……}，甚至整个集合N={0，1，2，3，4，……}或者空集φ={}。我们将只关心可计算性的问题，也就是："自然数的何种集合可由算法产生，何种不能？"

为了提出这样的问题，如果愿意的话，我们可把每一单独的自然数n，在一特别的形式系统中，以特定的符号串来表示。按照系统中（"语法正确"地表达的）命题的某一字典顺序，n表示"第n个"符号串，譬如讲Q_n 。则每一自然数代表一个命题。形式系统的所有命题的集合是由整个集合N来代表，例如，形式系统的定理可被认为是自然数的某一个更小的集合，例如集合P。然而，命题的任何特殊编号系统细节不是重要的。为了在自然数和命题之间建立一种对应，我们需要的是能从任一个自然数n得到它对应的（在一种适当的符号记法中写出的）命题Q_n 的已知算法，以及从Q_n 得到n的另一个已知算法。假定已知这两种算法，我们就能随心所欲地把一个特定形式系统的命题集合和自然数集合N相等同。

让我们选择一个形式系统，它是协调的，并广泛得足以包括所有图灵机的所有动作——并且在以下的意义上是"有意义的"，即它的公理和步骤法则可认为是"自明地真的"。现在，这形式系统的命题Q_0 ，Q_1 ，Q_2 ，Q_3 ，……中的一些实际上在该系统中有证明。这些"可证明的"命题有一些属于N的某一个子集的数，这事实上就是上面考虑的定理的集P。我们事实上已经看到了在某一个给定形式系统中存在一种一个接一

个产生具有证明的所有命题的算法。（正如早先概述的，“第n个证明”$\prod_n$ 是在算法上从n得到的。所有我们要做的是去看第n个证明的最后一行，以发现在系统中可证明的第n个命题，也就是第n个“定理”。）这样，我们就有了一个接一个（也许会有重复——但这无所谓）产生P的元素的算法。

一个可用某种算法以这种方式产生的集合，譬如P，叫作递归可数的。注意，在系统中可被证伪——也就是其否定的命题可被证明的命题的集合也类似地为递归可数的，因为我们可简单地列举这些可证明的命题。在此过程中取它们的否定。存在许多N的其他递归可数的集，但我们不必介绍把它们定义出来的形式系统。递归可数集的简单例子是偶数

{0，2，4，6，8，……}，

和平方的集合

{0，1，4，9，16，……}，

以及素数的集合

{2，3，5，7，11，……}.

很清楚，我们可以利用算法把这些集中的每一个元素产生出来。在这3个例子中还有这种情形，即集合的补集——也就是不在该集中的自然数的集为递归可数的。3种情形的补集分别为

{1，3，5，7，9，……}，

{2，3，5，6，7，8，10，……}，

以及

{0，1，4，6，8，9，10，12}.

为这些补集提供算法是轻而易举的事。我们的确可以在算法上决定，对于给定的自然数n，它是否为偶数，是否为平方或者是否为素数。这就为我们提供了既产生集合又产生补集的算法，因为我们可以按

顺序地跑过自然数，并在每种情况下决定它是否属于原先的集合或它的补集。一个本身及其补集都是递归可数的集合称为递归集。很清楚递归集的补集仍为递归集。

现在，是否存在递归可数但不是递归的集合呢？我们暂停一下，注意一下它的推论。由于这种集合的元素可在算法上产生，我们就有一种对于怀疑属于该集合的元素决定其是否真的属于该集合的手段。这一时刻，我们暂且假定它实际上属于该集合。所有我们要做的是允许我们的算法跑过集合中的所有元素，直到它最终找到我们所考察的特殊的元素。但是假如我们怀疑的元素实际上不在这集合中，则我们的算法就无济于事了。由于它会不断地进行下去，永远得不出一个决断。在这种情形下，我们需要一个产生补集的算法。如果它发现了我们所怀疑的，则我们肯定地知道该元素不在这集合中。我们用两种算法就应该是万无一失了。我们可以简单地交替使用这两种算法，并用任何一种方法找到所怀疑的。然而，这种快乐的情形只发生在递归集的情形下。我们这里只假定集合为递归可数的而不是递归的：我们提议的产生补集的算法不存在！这样，我们就面临着这等古怪的情形，即对于在集合中的一个元素，我们可在算法上决定它的确是在这集合中，但是我们用任何算法都不能保证决定恰巧不在这集合中的元素的这一个问题！

这种古怪的情形是否发生过呢？也就是说，是否的确存在不是递归的递归可数集呢？关于集合P的情况如何呢？它是一个递归集吗？我们已知它是递归可数的，所以我们必须决定其补集是否也为递归可数的。事实上它不是！我们何以知道呢？我们知道图灵机的动作被假定为在我们形式系统中允许的运算。我们用T_n来标志第n台图灵机，则陈述

“T_n（n）停止”

是一道命题——让我把它写作S（n）——也就是对于每一自然数n，我们可在我们的形式系统中把它表达出来。对于某些n值命题S（n）是真的，对于另外的n值它是假的。n跑过自然数0，1，2，3，……时所有S（n）的集合将由N的某一个子集S所代表。现在回忆一下图灵的基本结果（参阅第2章78页），在T（n）事实上不停的情形下，不存n在作“T_n （n）不停”断言的算法。这表明假的S（n）的集合不是递归可数的。

我们观察到S在P中的部分刚好包括了那些是真的S（n）。为什么会这样子呢？如果任何特别的S（n）是可证明的，那么它必须是真的（因为我们已选择了“有意义的”形式系统），所以S在P中的部分必须只包括真的命题S，而且没有真的命题S（n）能处在P的外头，因为如果T_n（n）停止，那我们便可在这系统内提供证明说它是真的这样[26]。

现在，假定P的补集是递归可数的。那我们就应有某种产生这种补集的算法。我们可以使这些算法运行并在其经过每一命题S（n）时记下来。这些都是错的S（n），所以我们的步骤实际上为我们递归地列举了错的S（n）的集合。但是，我们在上面注意到错的S（n）不是递归可数的。这一矛盾显示了，P的补集根本不是递归可数的；所以集P不是递归的，这就是我们所需要的结果。

这些性质在实际上表明了我们的形式系统不能是完备的，也就是说，在系统中必有一些既不能证明又不能证伪的命题。因为如果没有这样“不可决定的”命题，则集P的补集就必须为可证伪的命题（任何不能证明的东西都必须为可证伪的）。但是，我们已看到可证伪的命题包含一个递归可数集，所以这就使得P成为递归的。然而，P不是递归的，这一个矛盾导致了不完备性。这就是哥德尔定理的主要突破。

现在关于N中的代表我们形式系统的真的命题的子集T能说些什么呢？T是递归的吗？T是递归可数的吗？T的补集是递归可数的吗？事实上对所有这些问题的答案都是“否”。一种看到这一点的方法是注意到形式

“T_n（n）停止”

的假的命题不能由算法产生，正如我们前面所注意到的。所以，假的命题作为整体来说不能由任何算法产生，因为任何这种算法特别会列举出上面所有假的“T_n（n）停止”的命题。类似地，不能由一个算法产生所有真的命题（由于可轻易地修改任何这种算法以得到所有错误的命题，只要简单地把它产生的每一命题都取一个否命题即可）。由于真的命题因此不是递归可数的（假的也不能），它们构成了比系统中可证明的命题更复杂和深广得多的陈列。这再一次阐明了哥德尔定理的结论：形式论证只是得到数学真理的部分手段。

存在一定的真的算术命题的简单的族，却的的确确能形成递归可数集。例如，不难看出，具有如下形式的真的命题

$$\exists w, x\ldots\ldots, z[f(w, x\ldots\ldots, z)=0]$$

组成递归可数集（我把它记作A）[8]。这儿f（）是由通常的加、减、乘、除和自乘等算术运算所构成的。这种形式命题的一例——虽然我们不知它是否真的——是“费马大定理”的否定，此处f可取作

$$f(w, x, y, z) = (x+1)^{w+3} + (y+1)^{w+3} + (z+1)^{w+3}.$$

然而，人们发现集合A不是递归的（这是不容易看到的事实——虽然它是哥德尔原先论证的一个推论）。这样，我们并没有任何算法手段哪怕在原则上决定“费马大定理”的真假！

我试图在图4.1中极其概略地把所有具有好的简单的边界的区域代表一个递归集合，这样人们可以想象，告知某一给定的点是否属于该集是件直截了当的事。图中的每一点都认为代表一个自然数。而其补集也为一个显得简单的区域所代表。我在图4.2中试图用具有复杂边界的集合来代表递归可数但非递归的集合。此处边界一边的集合——递归可数的那一边——被认为比另一边简单。这些图是非常概略的，一点也没有在任何意义上的“几何准确性”的企图。尤其是用平坦的二维平面来代表这些图像在实际上没有任何意义。

我在图4.3中概略地指出了区域P, T和A处在集合N中的情形。

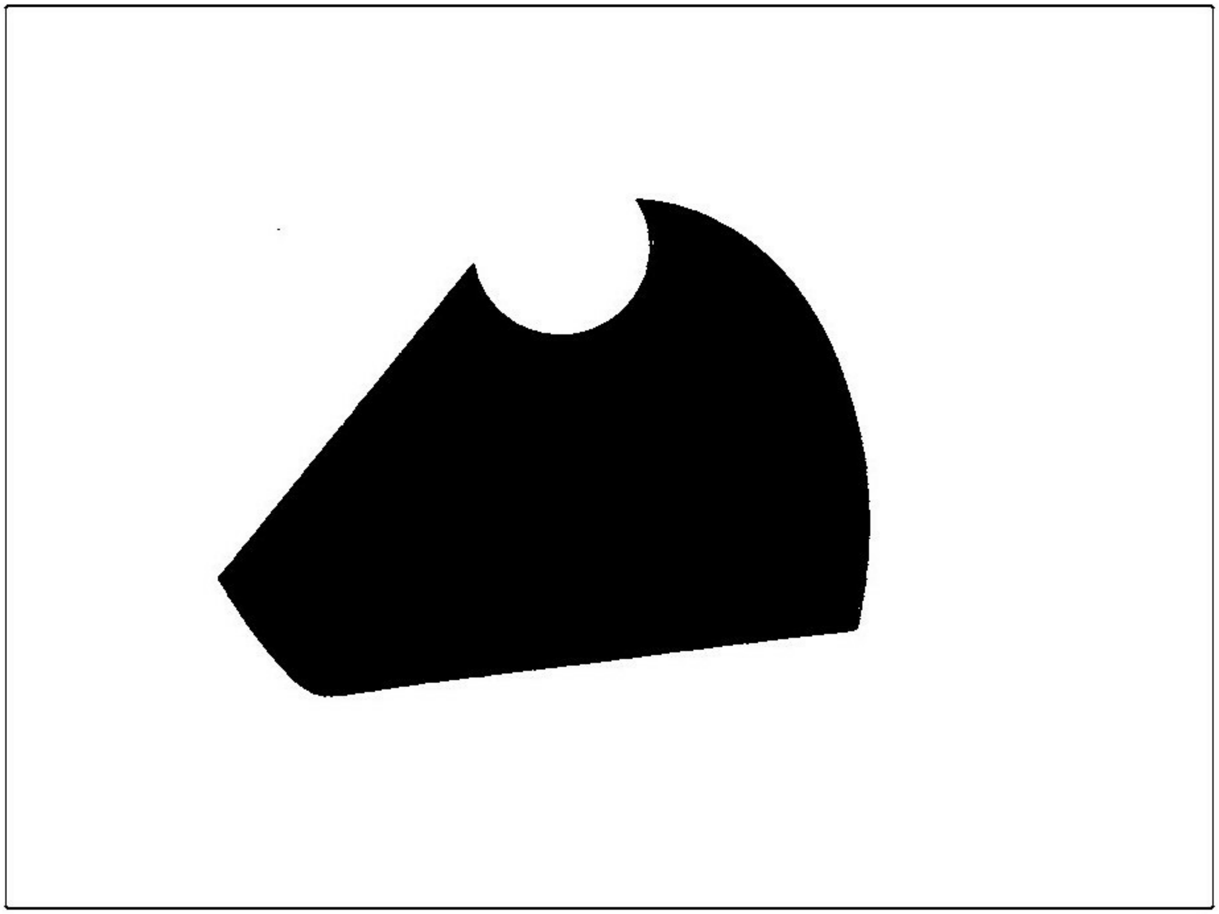


图4.1　一个递归集的高度概略的图示

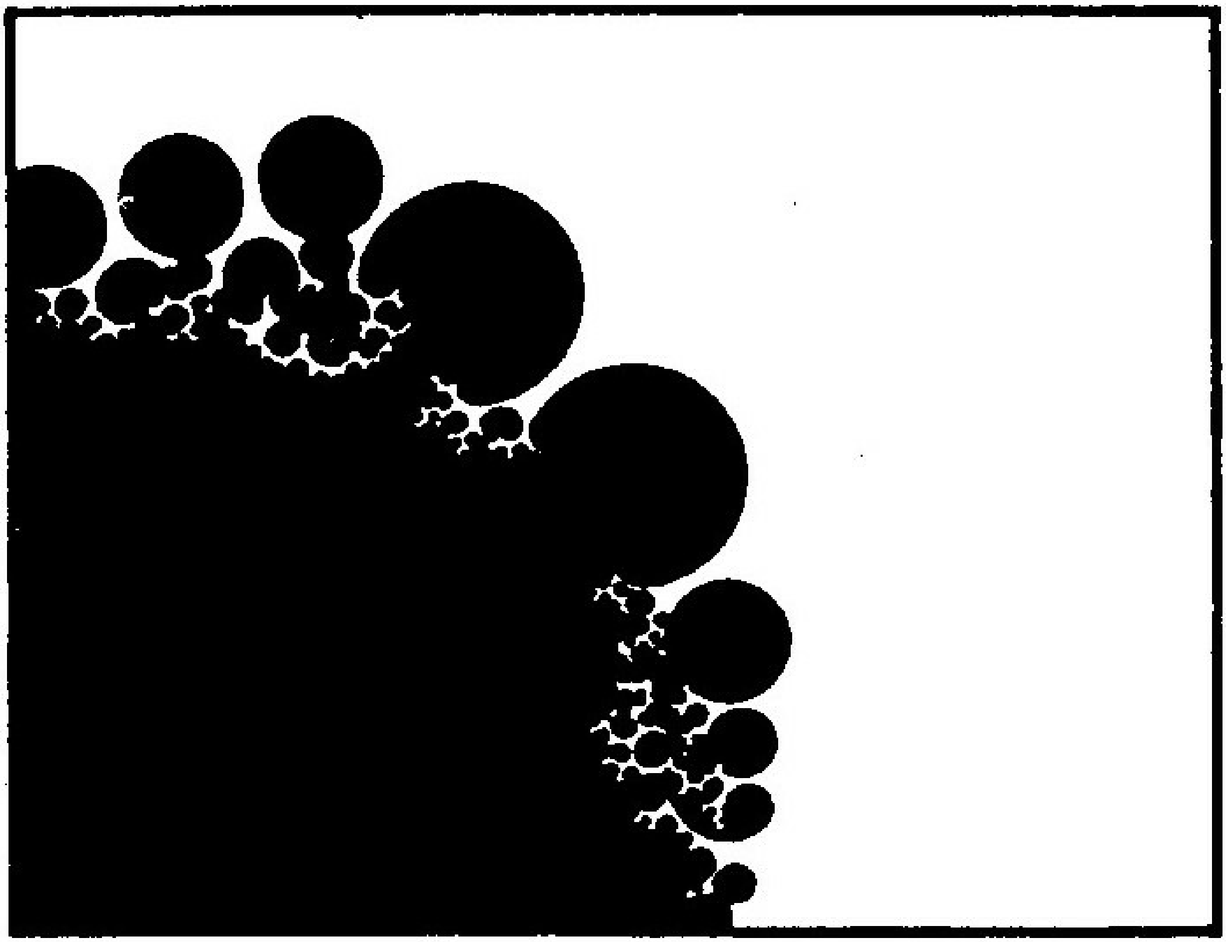


图4.2　一个递归可数的，但不是递归的集合（黑区域）的高度概略的图示。其思想是，白的区域定义为当可计算地产生的黑的区域被取走后所“余下的”；断定一点是否在白的区域中不是一个可计算的问题

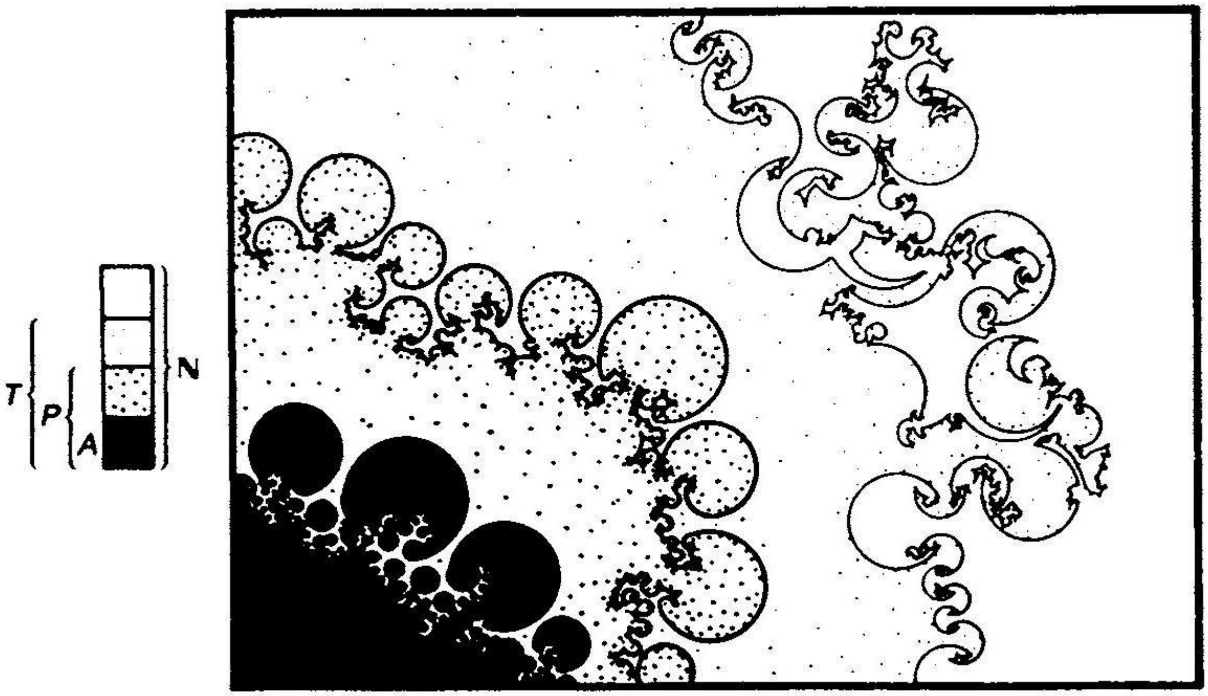


图4.3　不同命题集合的高度概略的图示。在系统中可证明的命题集合P，正如集合A那样，是递归可数但不是递归的；真的命题集合T甚至不是递归可数的

芒德布罗集是递归的吗

非递归集必须具有这样的性质，即它们在非常本质的方式上是复杂的。在某种意义上看，它们的复杂性应当公然抵抗任何系统化的企图，否则该系统化就会导致某种适当的算法步骤。对于一个非递归的集合，不存在一般的算法的方式去决定一个元素（或一“点”）是否属于这个集合。我们在第3章的开头肯定是见证到一个非同寻常地复杂的集合，也就是芒德布罗集。虽然提供其定义的规则是令人吃惊地简单，但集合本身却呈现出高度繁复的结构和无穷的变化。这难道真的是呈现在我们眼前的非递归集合的例子？

然而，读者会很快地指出，现代高速电脑的魔术把这些模式的复杂性呈现于我们的面前。难道电脑不就是算法行为的体现吗？的确，这肯定是对的。但是，我们必须记住电脑实际上产生此图的方式。为了检验亚根平面上的一点——一个复数C—是否属于芒德布罗集（涂成黑色）或它的补集（涂成白色），电脑就要从0开始，然后利用

$z \rightarrow z^2 + C$

把0映射到C，然后从z=C得到C^2+C，然后从$z=C^2+C$得到$C^4+2C^3+C^2+C$等。如果序列0，C, C^2+C, $C^4+2C^3+C^2+C$，……保持有界，则由C代表的点就涂成黑色；否则涂成白色。机器如何告知我们说这样的序列保持有界呢？这个问题原则上牵涉到知道在序列的无穷项后会发生什么，这本身不是电脑的事体。幸运的是，若序列是无界的，总存在有限项后就使人们得知的方法。（事实上，只要它达到以原点为中心以$1+\sqrt{2}$为半径的圆周就能肯定该序列是无界的。）

这样，在一定的意义上讲，芒德布罗集的补集（也就是白的区域）是递归可数的。如果复数C在白的区域中，就有确定此事实的算法。芒德布罗集本身也就是黑的区域的情况又如何呢？是否有确切告知一个被怀疑处于黑区域的点果真是在黑区域的算法呢？迄今看来这一问题的答案仍是未知的[9]。我询问了许多同事和专家，似乎没有人知道存在这样的算法。他们也从未表明过不存在这样的算法。对于黑区域至少还没有已知的算法。芒德布罗集的补集也许真正是一个递归可数但不是递归的集合！

在进一步探索这个设想之前，必须先讨论我掩饰的某些问题。这些问题对于以后讨论物理的可计算性具有某种重要性。我前面的讨论实际上有些不精确。我把诸如“递归可数的”和“递归的”这样的术语应用于亚根平面也就是复数的集合上。严格地讲，这些术语只能适用于自然数或其他可数的集合。我们已经在第3章（110页）看到实数是不可数的，所以复数也不是可数的——由于实数可考虑作特殊种类的复数，也就是虚部为零的复数（参阅114页）。事实上，刚好存在和实数“一样多”数目的复数，也就是C那么多。（粗略地讲，为了建立复数和实数之间的一一对应，我们可以把每一复数的实虚部各作小数展开，然后将其交叉地塞到相应实数的奇数和偶数位上去：例如复数3.6781……+i512.975……对应于实数50132.6977851……）

逃避这个问题的一种办法是只管可计算的复数，我们在第3章看到，可计算的实数——并因此可计算的复数——的确是可数的。然而，这里有严重的困难：事实上不存在决定两个按照它们相应的算法给出的可计算数是否相等的一般算法！（我们可以算法地形成它们的差，但我们不能算法地决定这个差是否为0。想象两个分别产生0.99999……和

1.00000……的算法，我们也许永远不会知道这些9和0是否无限地继续下去，因此这两个数相等，或最终某些其他的数会出现，因此这两个数不等。）这样，我们也许永远不能知道这些数是否相等。其中的一个含义是，甚至对诸如亚根平面上的单位圆盘这么简单的集合（所有到原点的距离不大于一个单位的点的集合，也就是图4.4中的黑的区域）都没有决定复数是否实际上处于圆上的算法。当点处于圆盘的内部（或在外部）时不会引起这个问题，但点处于圆盘的边缘时，也就是在单位圆本身上时就有了问题。单位圆被认为是圆盘的部分。假定我们简单地给出产生某复数的实部和虚部的位数的算法。如果我们怀疑该复数实际上处于单位圆上，我们并不能肯定这个事实。不存在去决定可计算数

$$x^2 + y^2$$

是否实际上等于或不等于1的算法，也就是决定该可计算复数x+iy是否在单位圆上的判据。

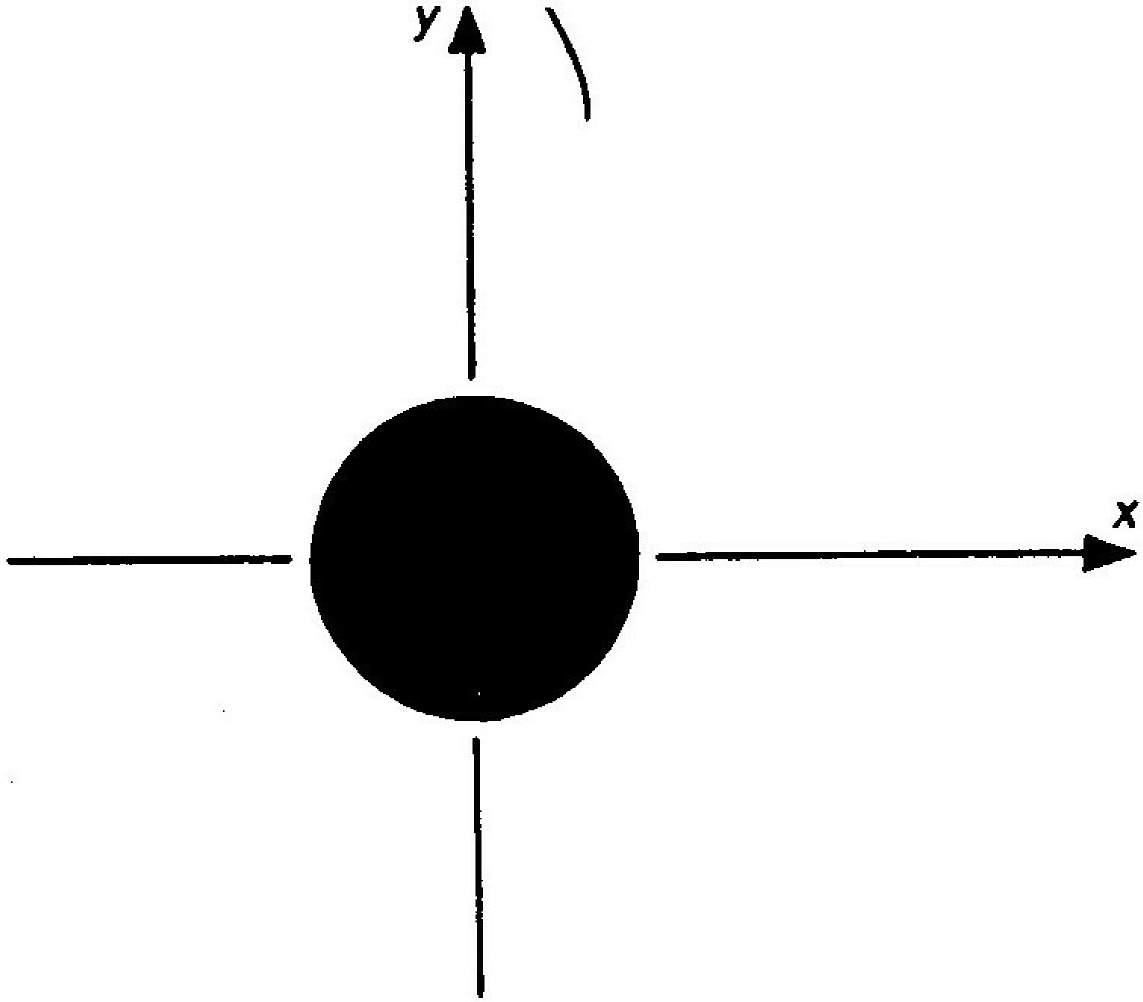


图4.4　单位圆盘肯定被当作“递归的”，但是这里需要一个适当的观点

这肯定不是我们所需要的。单位圆盘当然必须被当作递归的！没有很多集合比单位圆盘更简单！一种躲避这一问题的办法是不理睬边界。对实际上处于内部或外部的点肯定存在确认这些事实的算法。（简单地一个接一个地产生x^2+y^2的数位，最终会发现在小数展开0.99999……后面出现非9或1.00000……后面出现非0）。在这个意义上讲，单位圆盘是递归的。但是，这种观点是相当粗劣的，因为人们经常需要按照在边界上的行为来进行论证。另一方面，这种观点或许对物理学是合适的。我们以后还要再考虑这些问题。

人们或许还会采用另一种紧密相关的观点，它根本未涉及可计算复数的问题。我们简单地要求可对给定的复数决定其是否在该集中或在补集中的算法，而不试图去列举该问题的集外或集内的复数。我在这里

的“给定”的意义是，对于我们检验的每一个复数，也许用某种魔术的办法，实部和虚部的连续位数可一个接一个地写出以供使用，要多长就有多长。我不要求存在任何已知或未知的把这些位数写出来的算法。对于一个复数的集合，如果存在一个单独的算法，使得只要并且只要一个复数实际上在此集中，一旦该数以这种方法用一串数位写出，就在有限的步骤后它最终会说“是”，则该集合被认为是“递归可数的”。和上面提出的第一种观点一样，这种观点“不理睬”边界。这样，单位圆盘的内部和外部分别都在这个意义上被当作递归可数的，而边界本身不是。

我一点也不清楚，这些观点是否真正必需[10]。把它应用到芒德布罗集时，“不理睬边界”的哲学可能将该集合的许多复杂性都损失了。该集合一部分包括具有内部区域的“点”，还有部分是“卷须”。其极端复杂性似乎存在于极其剧烈的弯曲的卷须之中。然而，卷须不在集合的内部，所以如果我们采用了上述的任一种哲学，则这些卷须都被忽略了。尽管如此，当只考虑斑点时，仍然不清楚芒德布罗集是否为“递归的”。这个问题似乎依赖于某个未被证明的有关芒德布罗集的猜测：它是所谓的“局部连通”的吗？我不想在此解释此术语的意义及其关联之处。我只想指出这些是困难的问题，它们引起了有关芒德布罗集的未解决的问题，而且其中一些正是当前某些数学研究的最前沿的问题。

为了绕过复数是不可数的问题，人们还可以采用其他的观点。人们不去考虑所有可计算的复数，而去考虑这样的一个适当的子集，该子集的数具有去决定其中两个数相等与否仍是可计算的问题的性质。“有理”复数即为这样的一种简单的子集，实部和虚部均为有理数的复数即为有理复数。我认为它并不在芒德布罗集中占多少，而这种观点又是非常局限的。考虑代数数也许会更令人满意些——这就是那些为整系数的代数方程的解的那些复数，例如，方程

$$129z^7-33z^5+725z^4+16z^2-2z-3=0$$

所有z的解为代数数。代数数是可数的并且是可计算的。实际上去决定它们中的两个是否相等正是可计算的问题。（它们其中许多处于单位圆的边界和芒德布罗集的须蔓上。）如果需要的话，我们可把这问题表述成，芒德布罗集是否按照它们为递归的。

在刚才考虑的两个集合的情况下代数数也许是合适的，但它实在不能一般地解决我们所有的困难。考虑由关系

$$y \geq e^x$$

所定义的集合（图4.5中的黑的区域）。这里z=x+iy是亚根平面上的点。按照上面所表述的任何观点，该集合的内部及其补集的内部，都是递归可数的。但是（从F.林德曼在1882年证明的一个著名定理）边界 $y=e^x$ 只包含一个代数点，即z=i。代数数对于这种情形下的边界的算法性质的研究毫无用处！虽然不难找到满足这种特殊情形的其他的可计算数子集，但人们会强烈地感到，我们还没得到正确的观点。

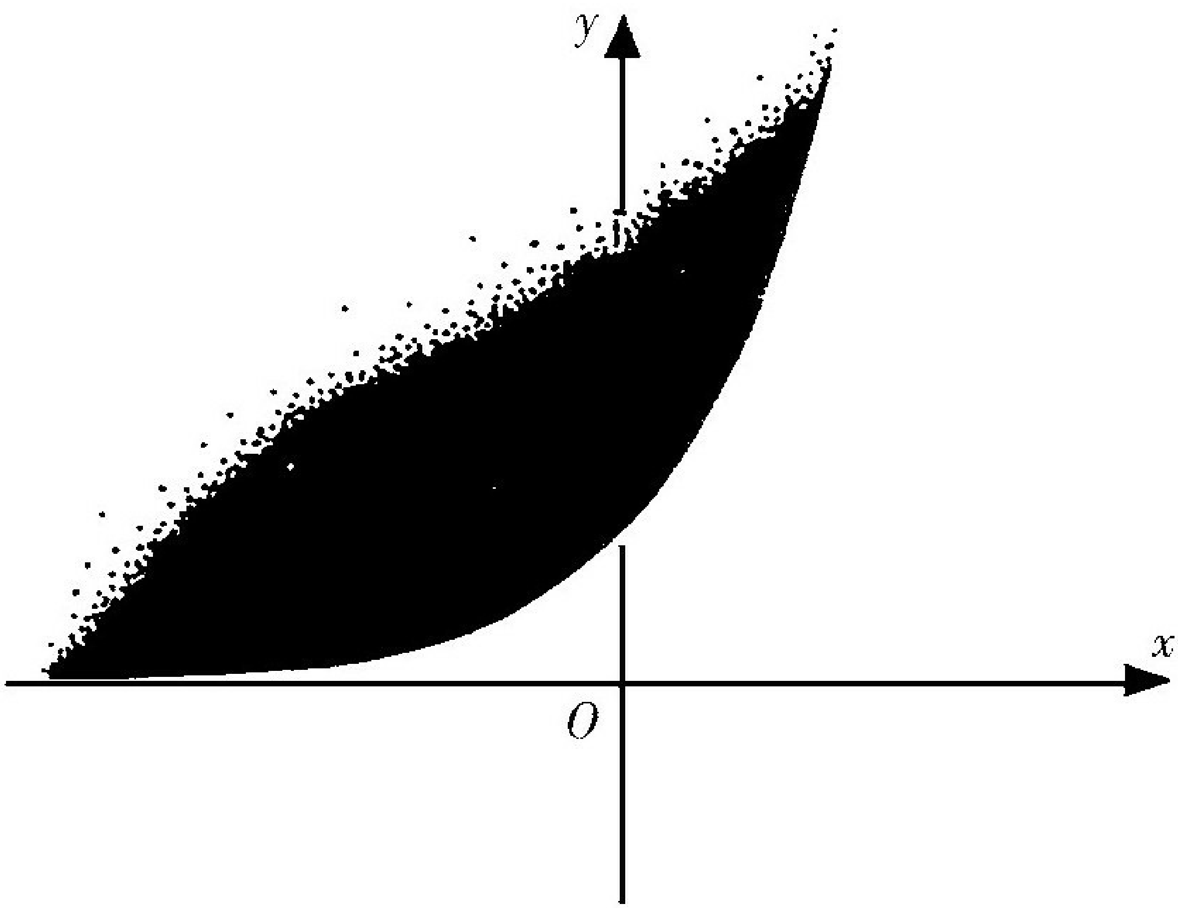


图4.5　由指数关系 $y>e^x$ 定义的集合也应当算作“递归的”

一些非递归数学的例子

在许多数学分支中产生了非递归的问题。也就是说，我们会遇到一系列的问题，它们答案或者为“是”或者为“非”，但是不存在决定究竟是什么答案的一般算法。在这类问题中有一些显得非常简单。

首先考虑求整系数代数方程组的整数解的问题。这种方程称为丢番图方程（以希腊数学家丢番图来命名，他的生活年代为公元前3世纪，他研究了这一类方程）。这样的一组方程可为

$z^3 -y-1=0$，$yz^2 -2x-2=0$，$y^2 -2xz+z+1=0$，

问题在于决定它们是否有x, y，z的整数值的解。在给定的特殊情况下，事实上存在

$x=13$，$y=7$，$z=2$

的解。然而，不存在决定任意丢番图方程集合[27]的这一问题的算法：尽管丢番图算术是这么初等，它却是非算法数学的一部分！

[另一个稍微高等的例子是流形的拓扑等价。这里我仅仅简略地提及，因为它和第8章要讨论的问题有某种可以预料到的相关性。为了理解何为“流形”，先考虑一个线圈，它是仅仅为一维的流形。然后考虑一个闭合面，这是二维的流形。再摹想具有三维或更高维的“表面”。两个流形的“拓扑等价”表明其中一个可以连续运动地变形成另一个——不能撕裂，也不能粘住。这样，一个球面和一个立方体的表面就是拓扑等价的，同时它们和一个环或茶杯的表面不是拓扑等价的——后两者实际上是相互拓扑等价的。现在，对于二维流形，存在一种决定其是否拓扑等价的算法——事实上可归结为计算每一曲面所具有“环柄”的数目。在写此书时，对于三维这问题的答案还没有得到，但是对于四维或更高维的情况，已经知道不存在决定等价类的算法。四维情形和物理有些相关是可以理解的。这是由于按照爱因斯坦的广义相对论，空间和时间一起组成了一个四维流形（见第5章268页）。格罗许和哈特尔在1987年提出，这个非算法性质可能和“量子引力”有关；还可参阅第8章。]

现在我们考虑一个被称作字问题[11]的不同种类的问题。假定我们有某些符号字母，考虑把这些符号连成各种称作词的串。词本身可以不具有意义，但是我们有一张（有限的）在它们之间“等价”的表，可用此

表来推导出更多这样的“等价”。这可以用如下办法做到，在较长的词中找出和表中某个词相同的部分，这一部分可用表中认为是相等的另一个词来取代。现在问题就归结为，对某一对给定的词，按照这些规则决定它们是否“相等”。

例如，我们原始的表为

EAT=AT

ATE=A

LATER=LOW

PAN=PILLOW

CARP=ME.

例如，从这些我们可以推出

LAP=LEAP

这可由连续地利用原表中的第二、第一以及再次利用第二个关系而得到：

LAP=LATEP=LEATEP=LEAP.

现在的问题在于，给定某一对词，我们能简单地用这种代入法从一个词得到另一个词吗？例如，我们能从ATERPILLAR得到MAN，或从CARPET得到MEAT吗？在第一种情形下的答案恰好为“是”，而在第二种情况下则为“非”。当答案为“是”时，通常显示这一点的方法是简单地写出一串等式，每一个词都是用允许的关系从前面的词得出。这样（要改变的字母用粗体印出，刚被置换的用斜体印出）：

CATERPILLAR=CAPRILLAR=CARPILLATER=CARPILLOW=CAR

按照允许的法则，我们何以得知不能从CARPET得到MEAT呢？对此问题，我们要稍微多想片刻，但是用各种不同的方法不难看到。最简单的方法如下：在我们原始表上的每个“等式”中，A加上W再加M出现

的总次数在两边是相等的。这样，在所有允许替代的系列中A、W和M的总数目不应改变。然而，对于CARPET这个数为1，而NEAT为2。所以靠允许的替换不可能从CARPET得到MEAT。

请注意，当两个词“相等”时，我们可简单地使用所给定的规则，写出一串允许的形式符号串来显示这一点；而在“不相等”的情形，我们必须求助于关于给定规则的论证。只要两个词事实上是“相等”的时候，我们就有清楚的算法可用来在它们之间建立起“相等”。我们所要做的是，把所有可能的词的序列作字典式的列表。如果序列中含有接连的两个词，其中第二个词不能按允许的规则从第一个词得出的，就从这表中删去这样的序列。余下的序列就提供了所有要寻找的词之间的“等价类”。然而，一般地不存在这样明显的算法，它能决定两个词不“相等”。为了建立这个事实，我们必须求助于“智慧”。（我的确花了好一阵时间才注意到上面的“技巧”，它可用来建立CARPET和MEAT的不“相等”。对于其他例子，也许需要完全不同的“技巧”。顺便提及，对于建立“等式”的存在，智慧虽然不是必要的，却是有助的。）

事实上，在上述情况中对于包含5个“等式”的特殊的表，当两个词的确“不等”时，提供一种去确定其“不等”的算法并不特别困难。但是，为了找到对这种情况起作用的算法，我们必须使用一些智慧！人们发现，并不存在任何单独算法可普遍地应用于所有原始表的选择。在这个意义上讲，字问题不存在算法解。一般字问题是属于非递归数学的范畴！

甚至对于某种特别选取的初始表，不存在决定两个词语何时不相等的算法。其中一例便是：

AH=HA

OH=HO

AT=TA

OT=TO

TAI=IT

HOI=IH

THAT=ITHT

（这是采用G.S.蔡亭和丹娜·斯各特1955年给出的表；参阅*Gardner 1958*，第144页。）这样，这个特殊的字问题本身就是一个非递归数学的例子。也就是说，利用这张特殊的初始表，我们不能在算法上决定两个给定的词是否“相等”。

从形式化数理逻辑的考虑（正如我们早先考虑过的“形式系统”等）中产生了一般的字问题。初始表起着公理系统的作用，词的替代规则起着步骤的形式法则的作用。从这种考虑引起了字问题的非递归性的证明。

作为非递归数学问题的最后一个例子，现在我们考虑一个用多边形来覆盖欧几里得平面的问题。这里我们只允许用有限种不同形状的花砖，看看是否能将整个平面既没有裂缝又没有重叠地覆盖住。这种用多边形来铺满平面的方法称为平面的镶嵌。我们都对如下事实很熟悉，可以只用正方形或正三角形或正六边形来镶嵌（正如第10章图10.2所示的），但是不能只用正五边形。还有许多其他的单独形状可以用来镶嵌平面，正如画在图4.6中的两种不规则五边形。用两个形状来镶嵌，结果就更精巧。图4.7画出了两个简单的例子。迄今为止所有的例子都具有称为周期性的性质。这表明它们在两个独立的方向上完全重复。按照数学语言，我们说存在一个周期平行四边形——一个平行四边形，如果我们用某种方法将其标出，并在平行于它的边的两个方向上不断地重复，则能重新产生给定的镶嵌花样。图4.8即为一个例子，在左面画出了用刺状的花砖进行的周期镶嵌，而在右面则画出与此周期性镶嵌相关的周期平行四边形。

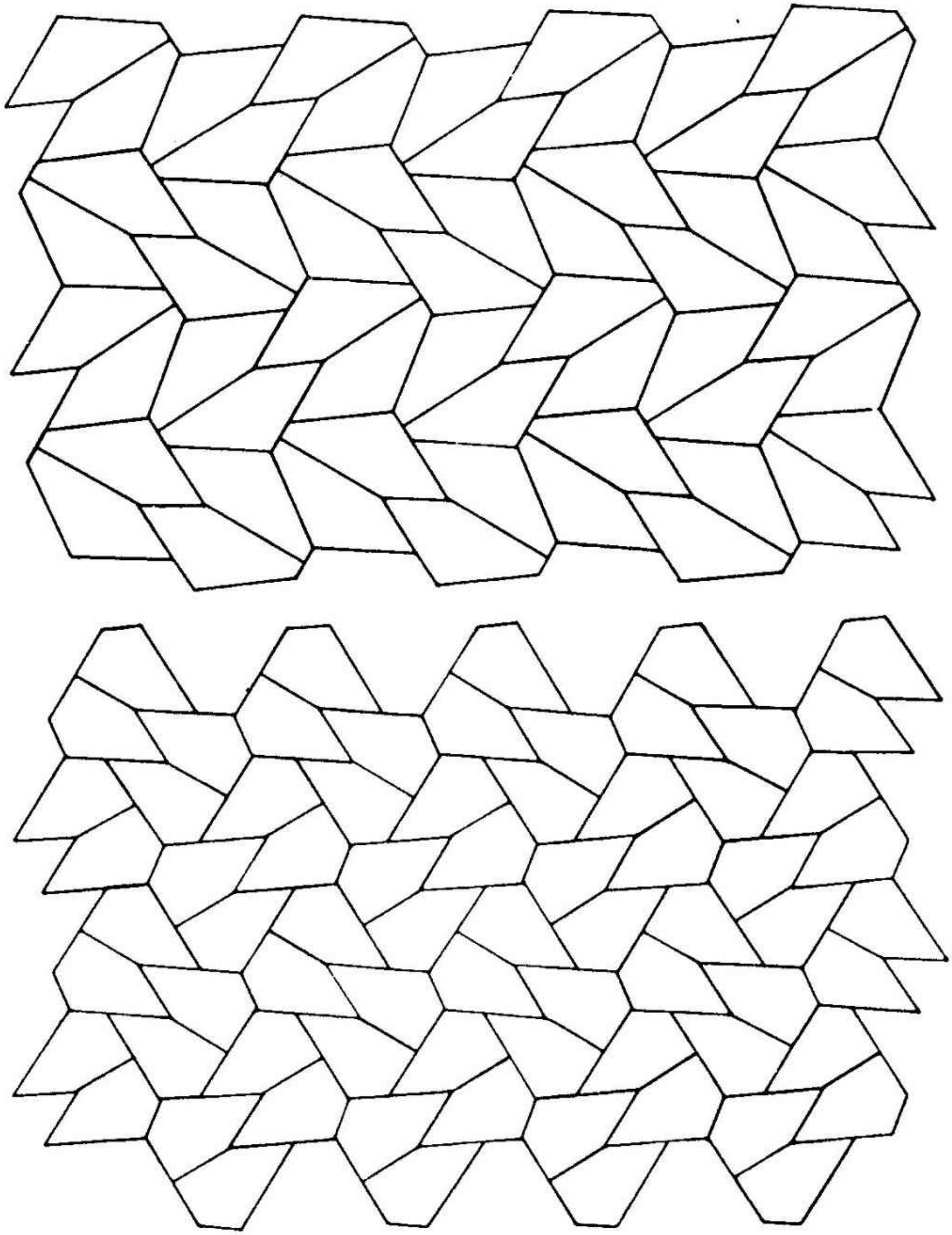


图4.6　平面周期镶嵌的两个例子，每一种情形都只用单独的花砖（1976年由马约里·赖斯发现）

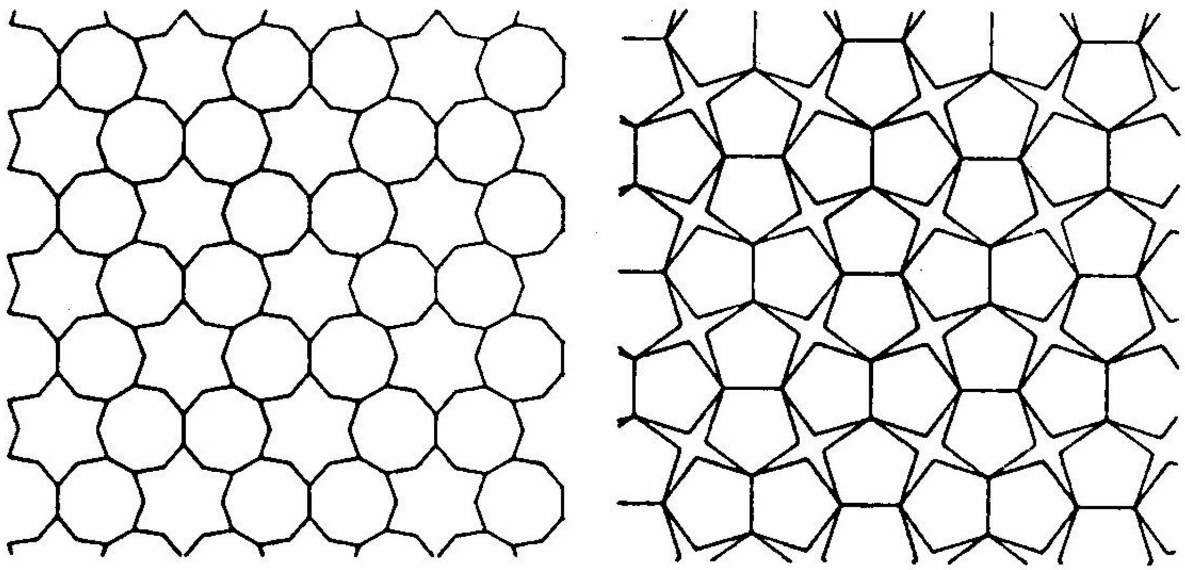


图4.7　平面周期镶嵌的两个例子，每一种情形都用两种花砖

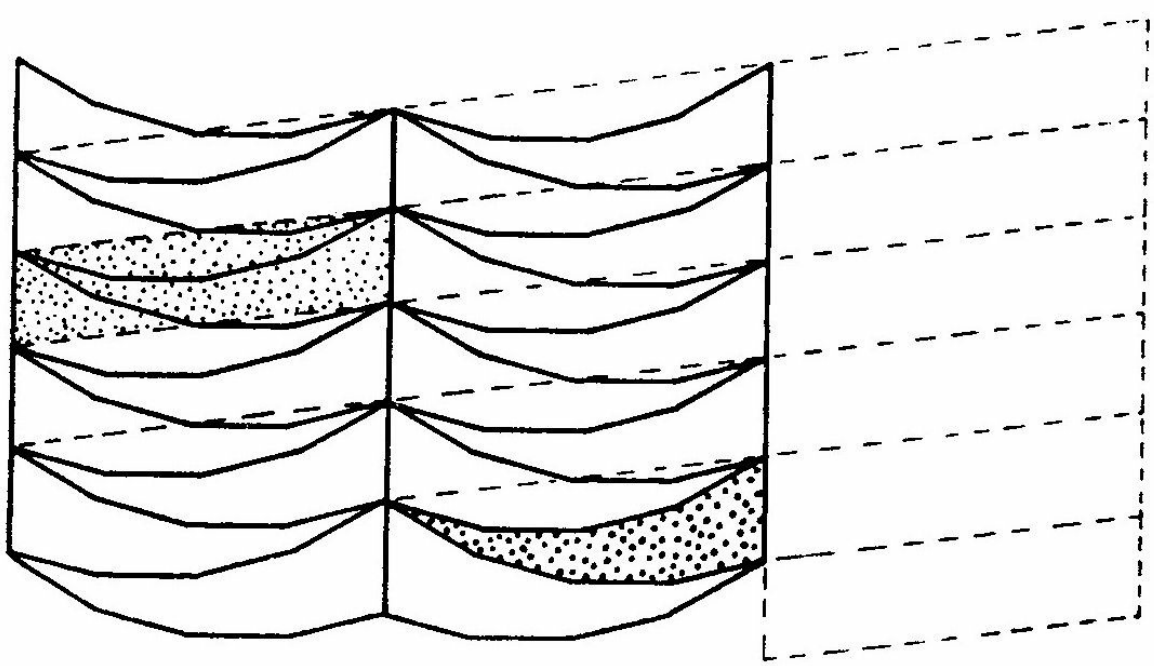


图4.8　一个周期性镶嵌，并标出和它的周期平行四边形的关系

存在许多不是周期性的平面镶嵌。图4.9画出了3种，这是用图4.8所示的同一种刺状花砖组成的非周期性的“螺旋”状镶嵌。这一种特别的花砖形状（由于明显的原因）被叫作“万能的”，它是由B.格吕堡和G.C.谢发德设计的（1981，1987），这明显的是基于H.冯德堡的更早的形状。

值得注意的是，用这种花砖既可以构成周期性的也可以构成非周期性的镶嵌。许多其他单独花砖形状和花砖集合也具有这种性质。现在我们要问，是否存在一种花砖或一组花砖，只能非周期性地镶嵌平面呢？答案是肯定的。在图4.10中我画出了一族由美国数学家拉飞逸·罗宾逊（1971）构作的6个花砖，它们只能够非周期性地镶嵌整个平面。

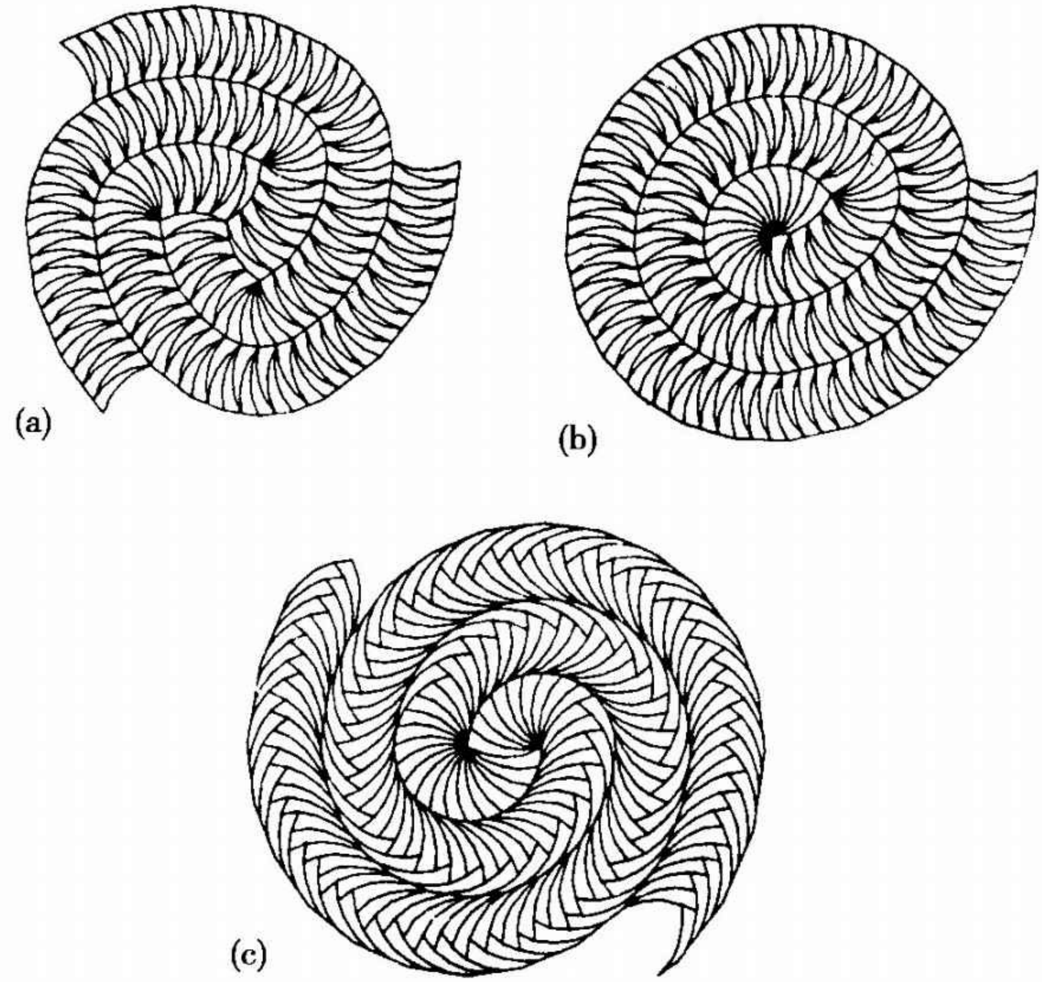


图4.9　3个非周期性的“螺旋”镶嵌，使用了图4.8中的同样的“万能”的形状

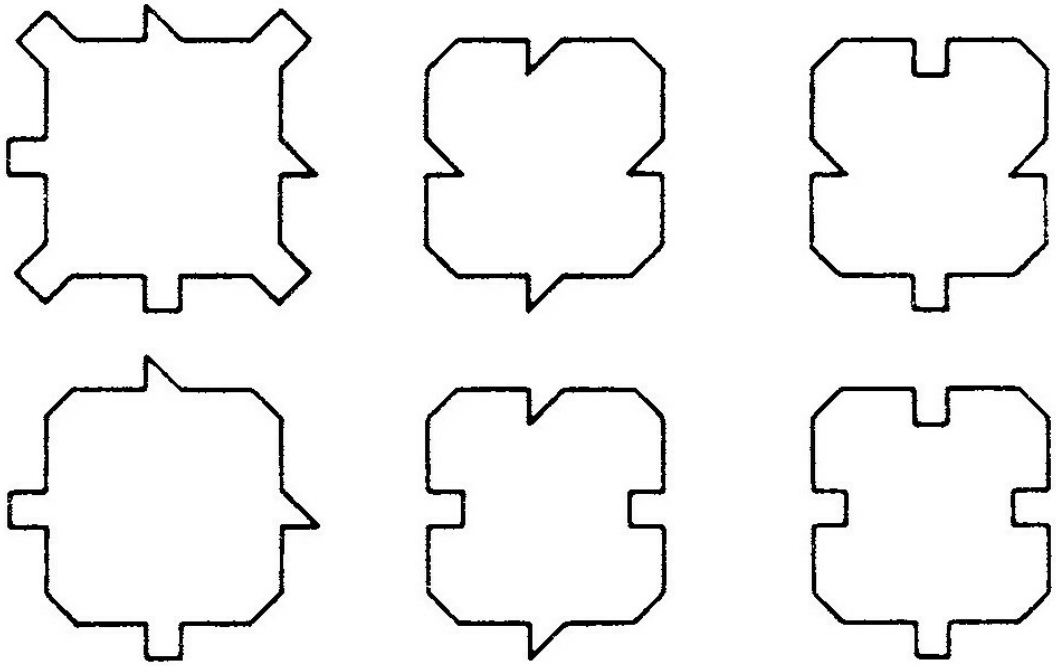


图4.10　拉飞逸·罗宾逊的只能对平面做非周期性镶嵌的6种花砖

值得稍微了解一下这种非周期性的花砖族由来的历史。（参阅*Grünbaum and Shephard 1987*）。1961年美籍华人逻辑学家王浩提出了对于镶嵌问题是否存在一个判定过程的问题，也就是说，是否存在一种算法，它可以判定给定的不同多边形的有限集合能否将整个平面镶嵌[28]！他指出，如果每一个以某种方式镶嵌平面的不同花砖的集合，还能把这平面周期性地镶嵌的话，则的确存在这样的决定步骤。我想，可能那时人们感到，不太会有违反这种条件的集合——亦即会存在“非周期性”的花砖集合。然而，1966年在王浩的建议指导下，罗伯特·伯格指出，镶嵌问题的判定过程实际上不存在；镶嵌问题也是非递归数学的一部分[12]！

这样，我们从王浩的早期结果得知，必然存在非周期性的花砖集合，而且伯格也确实找到了第一族非周期性花砖。但是，由于这些论证脉络之复杂性，他的集合涉及了非同小可的大数目的不同花砖——最初有20426个。伯特又用了许多技巧才将其数目减少到104个。然后到1971年，拉飞逸·罗宾逊将此数目减少到图4.10所示的6个。

图4.11中还画出了另外一种非周期性的6种花砖的集合。

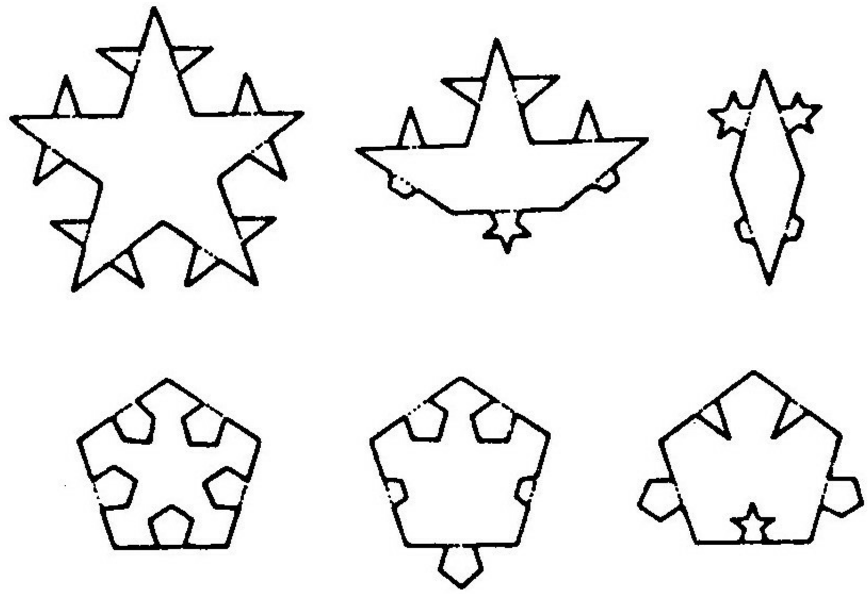


图4.11　另一种只能非周期性地对平面做镶嵌的6种花砖的集合

这是大约在1973年我自己沿着完全不同的思路得到的。（在第10章中我还将提及，图10.3画出了用这些形状铺就的排列。）我注意到罗宾逊的非周期性的6个集合后，开始设法减少此数目；试着拼拼凑凑，能够将其减少到两个。图4.12中画出了另外两种方案。这些完整的镶嵌显示出的必须为非周期性的花样，具有许多显著的性质，包括了似乎在结晶学上不可能的五重对称的准周期结构。以后我还会提及。

令人吃惊的是，数学中这么明显地“无聊的”领域——也就是用全等的形状去覆盖平面——初看起来像是“小孩游戏”，实际上应该是非递归数学的一部分。实际上，在这领域中还有许多未解决的困难问题。例如，我们还不知道，是否存在只包括单独花砖的非周期性集合。

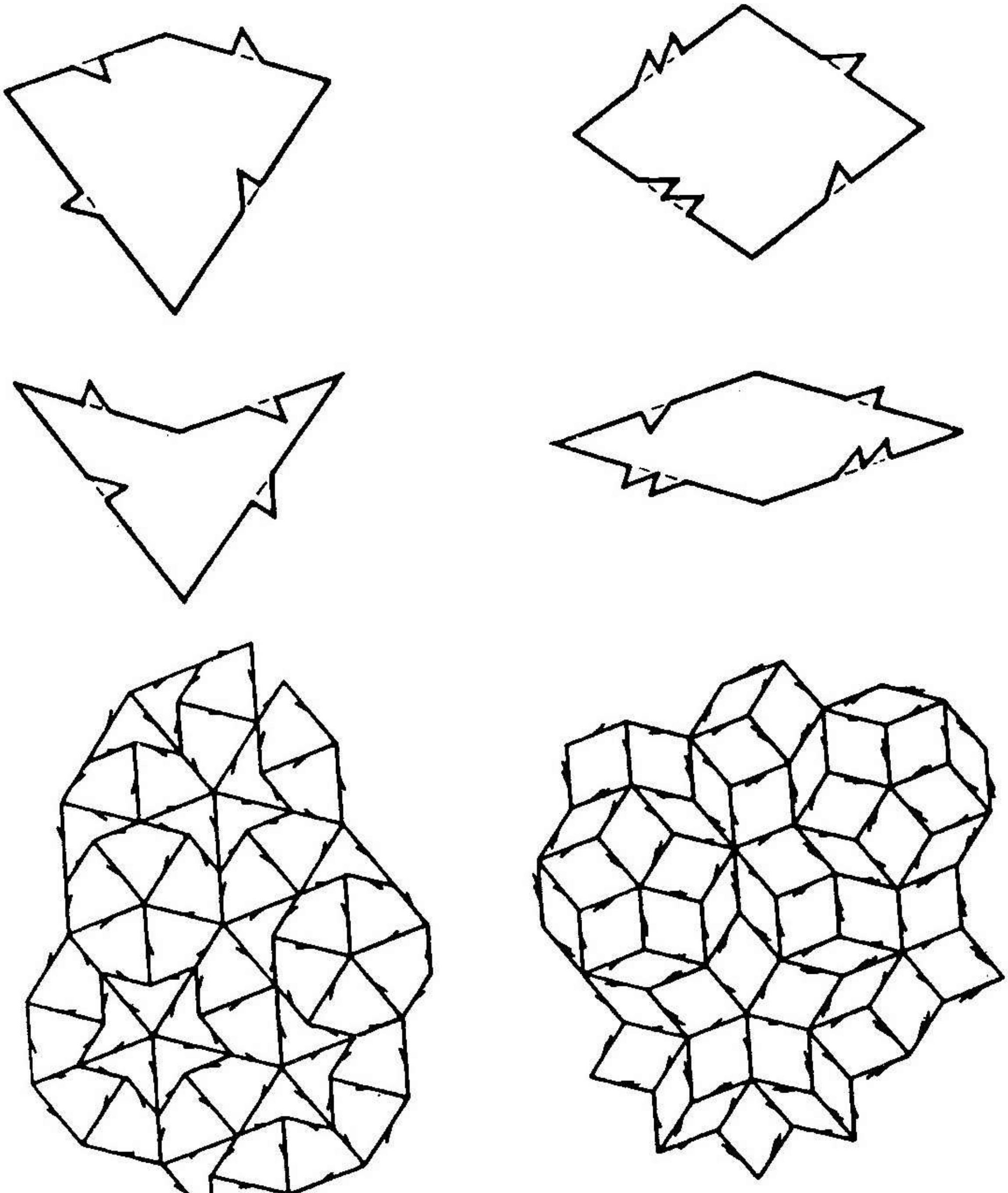


图4.12（彭罗斯镶嵌）两对花砖，每对都只能非周期性地镶嵌平面。还有由每对花砖镶嵌的平面区域

王浩、伯格和罗宾逊处理镶嵌问题时，所用的花砖是以方块为基础的。我这里允许用一般形状的多边形，而且为了展现独特的花砖，人们需要一些适合的可计算方法。一种方法是将其顶点当成复平面上的点，

也许这些点只要是代数数就完全足够了。

芒德布罗集像非递归数学吗

让我们回到早先的关于芒德布罗集的讨论。为了阐释的目的，我们假定，在某一适当的意义上，芒德布罗集是非递归的。由于它的补集是递归可数的，这就表明集合本身不是递归可数的。我认为，关于非递归集合和非递归数学方面，芒德布罗集的形式似乎对我们有许多教益。

回到第3章遇到的图3.2。我们注意到，集合的大部分似乎都由一个大的心状的区域所充满，在图4.13中用A来表示该区域。这个形状称为心脏线，它的内部区域可以定义为亚根平面的点c的集合。该集合是由

$$c=z-z^2$$

的形式产生的，z是离原点距离小于1/2的复数。这一集合在早先提出的意义上肯定是递归可数的：即存在一个算法，把它应用于区域的内部的一点时，将会断定这一点的确是在区域的内部。很容易从上述的公式得到实际的算法。

现在考虑刚好处于心脏线左边的圆盘状的区域（图4.13中的区域B）。它的内部区域为点

$$c=z-1$$

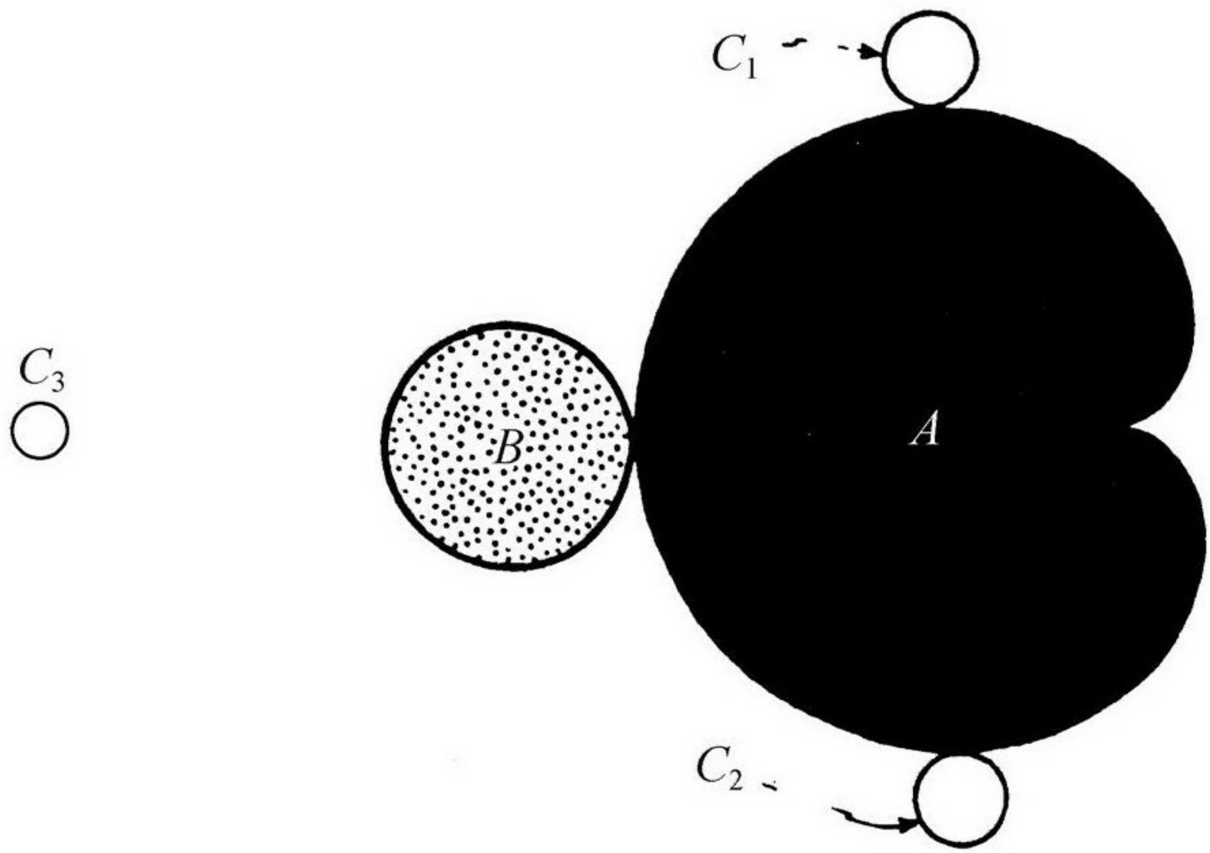


图4.13　可用简单的算法方程来定义芒德布罗集内部的主要部分

的集合，这里z离开原点距离小于1/4。这一区域的确是圆盘的内部——在一个正圆的内部的点集合。这一区域又是在上面意义下递归可数的。关于心脏线上其他“瘤”的情况又如何呢？考虑余下的两个最大的瘤。这是近似圆形的斑点，大致处于图3.2心脏线的上顶和底下，图4.13中用C_1，C_2 表示。它们可按

$$c^3+2c^2+(1-z)c+(1-z)^2=0$$

的集合给出，这里z的范围是离开原点距离小于1/8的区域。这个方程事实上不仅为我们提供了两个斑点（在一起的），而且还提供一个“婴儿”心脏线形状，后者出现在图3.2的左边的地方——也就是图3.1的主要区域——在图4.13中标作C_3 的区域。这些（一起或分开的）区域由于上述公式的存在又组成了递归可列集。

尽管我已经做过假设，即芒德布罗集可能是非递归的，我们运用某

些定义完好的以及不过于复杂的算法，可以清理出该集合的最大面积。这个步骤似乎应该继续下去。集合中所有最明显的、肯定占满了它面积的绝大部分（如果不是所有的话）的区域，可以在算法上处理。如果正如我所设想的那样，这集合全体不是递归的，则我们的算法不能达到的区域必须是非常精巧的，并且很难找到。而且，当我们已经定位了这样的一个区域，就可以看看有无机会改善我们的算法，使那些特殊的区域也能达到。然而（如果我关于非递归性的假设是正确的），还会有其他这类区域躲藏在微妙的、复杂的、模糊的深处，甚至用我们改善了的算法都达不到。我们再次可能利用直觉、天才和勤奋的巨大努力，将这样的一个区域定位，但是还会有其他的会漏掉，等等。

我想这就像用数学方式处理困难的问题，且假定为非递归性的。人们在某些特别的领域遇到的最普遍问题可由简单的算法步骤——甚至是已经知道了几世纪的步骤来解决。但是其中仍有漏网之鱼，要掌握它们就需要更复杂的步骤。漏网之鱼当然特别刺激数学家们，并促使他们去发展更为有力的方法。这些必须是基于对涉及的数学性质的越来越深刻的洞察之上。在我们对物理世界的理解中也许存在某些这种东西。

在上面考虑的字问题及镶嵌问题中，人们可以对这一类事稍有了解（虽然在这些领域中数学工具还未发展得非常远）。我们在一个非常特殊的情形下能用非常简单的论证去显示，某一词不能用允许的规则从另一词得到。不难想象，更复杂得多的推理可在处理更古怪的情形时起作用。很可能这些新的推理可发展成算法步骤。我们知道，不存在一个可以足够应付字问题的所有情况的步骤，但是漏掉的例子需要非常仔细和精巧地去构造。的确，只要我们肯定知道我们算法漏掉的例子，只要我们知道如何构造这些例子，则我们可以改善我们的算法以包括这种情形。只有不“相等”的配对词会漏掉，故一旦我们知道它们漏掉，我们就知道它们不“相等”，这一事实可添加到我们算法上去。我们改善了的洞察就导致一个改善了的算法！

复杂性理论

我在前面以及上一章关于算法的性质、存在和局限的论证是处于非常“原则的”水平上。我根本就没有讨论到出现的算法是否在任何方面像是可行的。即使对于算法存在并且该算法如何构造都很清楚的问题，也

还需要许多才干和勤勉，才能将此算法发展成有用的东西。有时小小的洞察和才干就能可观地降低算法的复杂性，以及有时极大地加快其速度。这些问题经常是非常精细和技术性的。近年人们在构造、理解和改善算法方面，在不同的情况下做了大量的工作。这是一个快速扩大和发展的研究领域。我不想对这些问题进行细致的讨论。然而，有关算法的速度可被增加的某一绝对的极限有各种普遍知道或猜测的东西。人们发现，甚至在具有算法性质的数学问题中，也存在种种内在的比其他问题更难于在算法上解决的问题。困难的问题只能用非常慢的算法（即，可能需要非同寻常地大量的存储空间的算法等）来解。有关这类问题的理论称为复杂性理论。

复杂性理论并不这么关心在算法上解决单个问题的困难，而是关心无限个问题的族，找到解决一个单独族的所有问题的一般算法。族中的不同问题会有不同的“尺度”。问题的尺度是由某一自然数n来测量。（关于这一个数n实际上如何表征问题的尺度，我一会儿还要再说。）算法对于每类中的每一特别问题所需的时间长度——或更正确地说，基本步骤的数目——是依赖于n的某一自然数N。稍微精确一点讲，我们讲在所有具有某一特别尺度n的问题中算法采用的最大的步骤数目为N。现在，当n变得越来越大，N也似乎变得越来越大。事实上，N似乎增加得比n快速得多。例如，N可以近似地和n^2，n^3 或许2^n 成比例（对于大的n，它比n, n^2，n^3，n^4 以及n^5 中的每一个都大多了，甚至比带有任何固定指数r的n^r 都大），或者譬如讲N甚至近似地和2^{2n} （这又更大得多）成比例。

当然，这些“步骤”的数目可依赖于实现该算法的电脑的类型。如果电脑为第2章描述的图灵机，那儿只有一盘磁带——这是相当低效率的——那数目N就会比允许两盘或三盘磁带的增加得更快速（也就是说机器会运行得更慢）。为了避免这类不定性，按照N作为n的函数增加的可能方式进行了宽广的分类，使得不管使用何种类型的图灵机，N的增加率的度量总是归到同一分类中去。一种称为P（说明“多项式时间”的分类包括了所有最多为n, n^2，n^3，n^4，n^5，……中的一个的固定倍数[29] 的速率。也就是说，对P分类中的任何问题（这里我的“问题”的真正含义是具有解决它们的一个一般算法的一族问题），我们有

$$N \le K \times n^r,$$

这里K和r为常数（与n无关）。这表明N不比n的某一固定方次的某一倍数更大。

两个数相乘肯定是属于P问题的简单类型。为了解释这一点，我必须首先描述数目n如何表征一对特殊乘数的尺度。我们可以想象每一个数都以二进制写出，而每一个数的二进制位数简单地为n/2，总共给出了n个二进制数——也就是总共n比特。（如果一个数比另一个短，可以简单地从短的开始连续地在前头加上零使之和长的具有一样的长度。）例如，如果n=14，我们可以考虑

1011010×0011011

（就是1011010×11011，但是在短的数上添了一些零）。最直接进行乘法的方式是只要写出：

```
              1 0 1 1 0 1 0
×             0 0 1 1 0 1 1
---------------------------
              1 0 1 1 0 1 0
            1 0 1 1 0 1 0
          0 0 0 0 0 0 0
        1 0 1 1 0 1 0
      1 0 1 1 0 1 0
    0 0 0 0 0 0 0
  0 0 0 0 0 0 0
---------------------------
  0 1 0 0 1 0 1 1 1 1 1 1 0
```

记住，在二进制中，0×0=0，0×1=0，1×0=0，1×1=1，0+0=0，0+1=1，1+0=1，1+1=10。单独二进制乘法的次数为（n/2）×（n/2）=n^2/4，并且可具有n^2 /4-（n/2）次的单独的二进制加法（包括移位）。这样，总共有（n^2 /2）-（n/2）次的单独算术运算——我们必须包括一些涉及移位的额外的逻辑步骤。总的步骤数为N=n^2/2（忽略低阶项），这

肯定是多项式的[13]。

一般来说，对于一族问题，我们取这问题的“尺度”的测度n为需要指明该特别尺度的问题的自由数据所需要的二进制位数（或比特）的总数。这意味着，对于给定的n，在给定的尺度下问题会有多到2^n 种不同的情形（因为每一位可有两种可能性中的任一个，0或1，而总共有n位数），而这些都必须由算法在不多于N步骤下被一致地处理好。

存在许多不属于P问题（的族）的例子。例如，为了进行从自然数r计算2^{2^r} 的运算，甚至只要写出这一答案就大约需要2^n 步骤，且不说进行计算了。n为在r的二进制表示中的位数。计算$2^{2^{2^r}}$ 的运算，只要写下就需要2^{2^r} 个步骤等！这些比多项式大多了，所以肯定不在P中！

在多项式时间内可以写下答案并甚至能检查正确与否的问题更为有趣。由此性质表征的（在算法上解出的）问题（的族）是一个重要的范畴。它们被称为NP问题（的族）。更精确地讲，如果在NP中的问题的族的个别问题有一解，那么该算法将给出这个解，并且它必须能在多项式时间内检验所设想的解确实是一个解。在问题没有解的情形下，算法会告诉我们这个，人们不必在多项式或别的时间内去检验的确没有解[14]。

NP问题既在数学本身，也在实在世界的许多范围内出现。我只给出一个简单的数学例子：在一个图中寻找所谓的“哈密顿回路”的问题（一个极简单概念的吓唬人的名字）。用“图”来表示点或“顶点”的有限集合，一定数目的点对由称为图的“边”的线连接起来。（我们在这儿并不对几何或“距离”性质感兴趣，只对由哪一顶点连接到哪一顶点感兴趣。这样，所有顶点是否在一个平面上表出是无关紧要的，我们的边是否互相穿越还是处于三维空间中都是无所谓的。）哈密顿回路就是一个只包括图的边的简单的闭合圈，该回路通过每一顶点刚好一次。图4.14中画出了一个在上面标出哈密顿回路的图。哈密顿回路问题是要判定，对于任何给定的图是否存在哈密顿回路，只要存在就把它明了地画出来。

可以以二进制数字用不同方式来表述一个图。用何种方法关系不大。一个步骤是给顶点编上号1，2，3，4，5，……，然后以某种适当的固定顺序列出成对的顶点来：

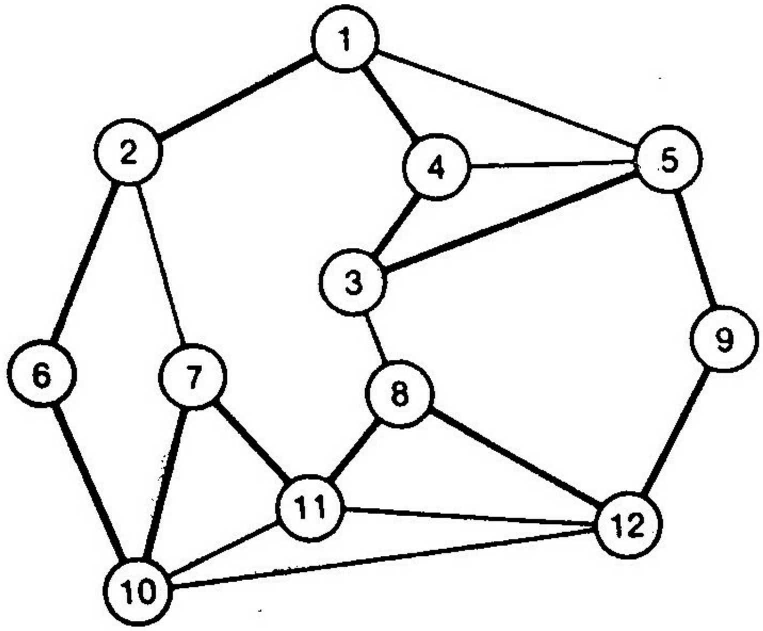


图4.14　带有（用稍粗一些的黑线）标出的哈密顿回路的一个图。还存在另一条哈密顿回路，读者若有兴趣可把它找出来

（1，2），（1，3），（2，3），（1，4），（2，4），（3，4），（1，5），（2，5），（3，5），（4，5），（1，6），……

然后我们做一个准确的0和1的搭配的表，当一对顶点对应于图的一个边缘时写上1，否则写0。这样二进制序列

100101101100……

表明顶点1接到顶点2、顶点4以及顶点5，……顶点3接到顶点4和顶点5，……顶点4接到顶点5，等等（图4.14）。如果需要的话，哈密顿回路可由这些边的子集给出，它用具有比前述的更多个零的二进制序列来描述。检验步骤是可以比开始找这些哈密顿回路更迅速地完成的事。人们所要做的一切，是检验作出的回路的确是一回路，也就是边须属于原先的图，而图中的每一顶点刚好只被用过两次——在两条边的每一端各一次。这一检验过程是某种可以在多项式时间内完成的事。

事实上，这个问题不仅是NP的，而且被认为是NP完备的。这表明其他任何NP问题都可在多项式时间内转变成它。这样，如果某个足够聪明的人能在多项式时间内找到解决哈密顿回路的算法，也就是能显示

哈密顿回路问题实际上是在P中！则其推论是任何NP问题都在P中！这样的事情具有重大的含义。一般地讲，对于适当大的n，在一台快速现代电脑上，P中的问题被认为是“易处理的”（也就是“在一可接受的”时间长度里是可解的）。而在NP中又不在P中的问题，对于相当大的n被认为是“不易处理的”（也就是虽然在原则上可解，

“在实际上是不可解的”），而不管我们将面临着何种可以预见的种类的电脑速度的增加。（对于大的n的不在P中的NP问题，需要的时间会急速地变得比宇宙的年龄还要长，这对于实际的问题没有什么用处！）任何在多项式时间内解决哈密顿回路问题的聪明算法都能转换成在多项式时间内解决任何其他NP问题的算法！

另一个NP完全问题 [15] 是“流动推销员问题”。这个问题和哈密顿回路问题很相像，只不过是在不同的边附上数字，人们寻求数（推销员走的“距离”）的和为极小的哈密顿回路。流动推销员问题的多项式时间解会又一次导致所有其他NP问题的多项式时间解。（如果真的找到这样的一个解，将会变成头条新闻！尤其是好几年来提出了密码系统，该问题有赖于大整数的因子化问题，这是另一种NP问题。如果可在多项式时间内解决这一问题，那么这样的码就可能被强大的现代电脑所破。但是如果不能，这码就是安全的。参见*Gardner 1989* 。）

专家们普遍相信，不管用任何种类图灵机的仪器，实际上都不可能在多项式时间内解决一个NP完全问题。所以结论是，P和NP不是同样的这个信念很可能是正确的，虽然还没有一个人能证明之。这仍然是复杂性理论最重要的未解决问题。

物理事物中的复杂性和可计算性

复杂性理论对于本书的讨论是重要的，因为它引起了另外的问题，和事物是否可计算的问题有一点区别；也就是说，被认为是算法的事物实际上是否以一种有用的方式为算法的。在后面的章节中，我关于复杂性理论将讲得比可计算性更少。因为我倾向于认为（虽然毫无疑问地是在相当不足够的基础上）复杂性理论的问题和可计算性本身的问题不一样，在和精神现象相关联上不占有中心的地位。而且，我感到算法的可行性问题的现状才刚刚被复杂性理论所触及。

然而，关于复杂性作用的问题，我也很可能是错的。正如我将在后面（第9章506页）评论的那样，实在物理对象的复杂性理论也许和我们刚刚讨论的有显著的不同。为了使这种差异变得更明了，那就必须使用量子力学，这个关于原子、分子的状态，以及许多其他在大得多的尺度下重要现象的神秘而强有力的精确理论。在第6章我们将遇到这一理论。按照最近大卫·德义奇（1985）提出的一系列思想，在原则上可能建造“量子电脑”，存在不属于P的然而由这种装置可在多项式时间内解的问题的（类）。直到现在一点也不清楚，一个实际的物理仪器如何建造成行为可靠的量子电脑，而且迄今所考虑的问题的（类）肯定是很人为的——但是，我们似乎已经知道了用量子物理仪器改善图灵机的理论可能性。

我在这里讨论时把它当作一台“物理仪器”的人类的头脑，尽管设计得非常微妙精巧，而且非常复杂，它本身会从量子理论的魔术中得到好处吗？我们是否理解量子效应可以用于解决问题或作判断的方式呢？为了利用这种潜在的好处，我们也许必须“超越”现存的量子理论，这是可以理喻的吗？实际物理仪器真的很可能改善图灵机的复杂性理论吗？实际物理仪器的可计算性理论又是如何呢？

为了研究这类问题，我们必须离开纯粹数学的领域，并在下面的章节中探求物理世界在实际上是如何行为的！

第5章　经典世界

物理理论的状况

为了了解意识为何是自然的一部分，我们对自然的运行要知道哪些呢？制约身体和头脑组成基元的定律与此关系重大吗？如果真的像许多人工智能的拥护者所竭力说服我们的那样，意识理解仅仅是由算法所规定的，那么这些定律实际上是什么样子的则是无关紧要的。任何能够实现算法的仪器都一样好。另一方面，也许我们的知觉比可怜的算法更富有内容。也许构成我们的确切方式正和实际上制约构成我们物质的物理定律一样重要。我们也许需要理解构成物体物质以及规定所有物体行为的根本性质。物理学尚未做到这一步。许许多多的秘密还有待揭示和探索。然而，大多数物理学家和生理学家却断言，我们已经拥有足够的关于通常尺度的、诸如人脑物体运作的物理定律的知识。作为一个物理系统，大脑毫无疑义是极端复杂的，我们对其结构的大部分细节和相应功能相当无知。几乎没有人说，人们对作为构成其行为基础的物理原则的理解不存在任何重大缺陷。

相反地，我在下面将持一种非传统的论点，也就是我们对物理学的理解，甚至在原则上还不足够用以描述我们大脑的运作。为了论证这一点，我首先必须概述物理学的现状。本章是关于所谓的“经典物理”，它包括牛顿力学和爱因斯坦相对论。此处“经典”基本上是指在1925年左右发现量子理论之前的占统治地位的理论。量子力学是由诸如普朗克、爱因斯坦、玻尔、海森伯、薛定谔、德布罗意、玻恩、约旦、泡利以及狄拉克的开创性工作的成果。它是一种不确定的、非决定性的、神秘的，描述分子、原子和亚原子粒子行为的理论。相反地，经典理论是决定性的，这样，将来总是由过去所完全固定。尽管许多世纪以来对经典物理学的理解使我们得到了非常精确的图像，它仍有许多神秘之处。我们还必须考察量子理论（在第6章）。因为和大多数生理学家的观点相反，我相信量子现象似乎对大脑的运行是相当重要的——这些是下面几章的内容。

迄今为止科学已取得了引人注目的成就。我们只要环视四周即可见

证理解自然帮助我们取得了何等伟大的威力。现代世界的技术大多是从大量的经验中推导出来的。然而，正是物理学理论以更基本得多的形式成为我们技术的基础。这正是我们在此所关心的。我们的理论是相当精确的。但其力量并不仅仅在此，而且在于异乎寻常地遵从精密的、微妙的数学处理的这个事实。正是这两者一道为我们带来了威力无比的科学。

这个物理理论的大部分并不特别新颖。如果首先要挑选一个事件的话，那应该是1687年艾萨克·牛顿出版了《原理》一书。这本重要著作向人们展示了如何仅仅从几个基本的物理原理出发，能够理解并经常以惊人的精度预言了大量的物理对象的行为。《原理》一书中很大部分是关于数学技巧的非凡的发展，尽管欧拉等人后来提供了实用的方法。正如牛顿所坦率承认的，他自己的工作大大得益于更早期思想家的成果，其中最杰出者为伽利雷·伽利略、雷奈·笛卡儿以及约翰斯·开普勒。还用了一些更古老的思想家们所奠定的重要概念，诸如柏拉图、欧多克索斯、欧几里得、阿基米德以及阿波罗尼奥斯等人的几何概念。我在下面还要更多地说到这些。

后来出现了对牛顿动力学基本框架的偏离。首先是19世纪中叶由詹姆斯·麦克斯韦发展的电磁理论。这个理论不仅包括了电场和磁场，而且还描述了光的经典行为 [1] ！此一杰出的理论将是本章后面所关注的课题。麦克斯韦理论对于今天的技术具有相当的重要性，并且毫无疑义地，电磁现象和我们大脑的工作密切相关。然而，和阿尔伯特·爱因斯坦名字联结的两种伟大的相对论对我们的思维过程是否具有任何意义，还没有这么清楚。亨利·庞加莱、亨德里克·安东·洛伦兹以及爱因斯坦为了解释当物体以接近于光速运动时所产生的使人迷惑的行为，从研究麦克斯韦方程出发，提出了狭义相对论（后来赫曼·闵可夫斯基给出了精巧的几何描述）。爱因斯坦著名的$E=mc^2$ 方程是该理论的一部分。但是迄今为止此理论对技术的影响（除了对核物理的效应之外）甚微，看来它和我们头脑工作的关联最多也只是外围的。另一方面，狭义相对论加深了我们对和时间本质有关的物理实体的理解。我们将会在后面几章看到，这给量子理论带来一些根本的迷惑，这些迷惑和我们对“时间流逝”的感觉有重要关系。况且，人们在鉴赏爱因斯坦的广义相对论之前必须理解狭义相对论。广义相对论是用弯曲的空间——时间来描述引力。迄今为止此理论对技术的效用几乎是不存在的 [30] ，看来极端地假设其对我们头脑的功能有何相关真有点异想天开了！然而，值得注意的

是，广义相对论的确和我们后面特别是在第7章和第8章的思考关系重大。在那里为了探索要获得量子理论首尾一贯的图像所必需的一些变动，我们要最彻底地研究空间和时间，—这些在后面还要更仔细地讲到！

经典物理学的领域很广阔。量子物理学的情况又如何呢？和相对论不同的是，量子理论正开始剧烈地影响技术。其部分原因在于，它为某些技术上诸如化学和冶金等重要领域提供了理解。人们的确可以讲，正是因为量子理论赋予我们新的详细的洞察力，才使这些领域被包含在物理之中。此外，量子理论还提供了许多全新的现象，我想最熟知的例子便是激光。量子理论的某些基本方面会不会在我们的思维过程的物理学中起关键的作用呢？

我们关于更现代的物理学能说些什么呢？一些读者也许会想起那些激动人心的观念，包括诸如“夸克”（参阅200页）、“GUT”（大统一理论）、暴胀宇宙论（参阅587页的注释13）、“超对称”、“（超）弦理论”，等等。将这些方案和我刚才提到的那些相比较又如何呢？我们是否也必须通晓这些呢？我相信，为了更清楚地透视，可将基本的物理理论分成三大类。我将这三类命名为：

1.超等的；

2.有用的；

3.尝试的。

本段之前所讨论的一切理论都必须归于超等类中。我并不强求只有该理论无可辩驳地适用于世界上的一切现象时才能称为超等的。但是，我要求在适当的意义上，该理论适用的范围和精确度必须是显著的。就我所理解的“超等”这个术语而言，居然会有属于这一类的理论存在，这真是极其令人惊异的事！我不知在其他科学中是否有理论可以归入这一类。也许达尔文和华莱士提出的自然选择庶乎近之，但还差得相当远。

我们在中学学到的欧几里得几何是一种最古老的超等的理论。古代人也许根本不将其当作一种物理理论，但实际上它的确是物理空间以及刚体几何的卓越的理论。为何我将欧几里得几何归于物理理论而不是数学的一个分支呢？具有讽刺意义的是，现在我们知道，欧几里得几何不

能当作我们实际生活其间的物理空间完全准确的描述，而这是采取这个观点的一个最清楚的原因！爱因斯坦的广义相对论告知我们，在引力场存在时，空间（—时间）实际上是“弯曲的”（也就是说不是完全欧几里得型的）。但是这个事实并没损坏欧几里得几何的超等的资格。在一米的尺度上，与欧几里得的平坦性偏差的确非常微小，它比一个氢原子的直径还小！

阿基米德、帕波斯和斯蒂文研究静态物体，并将其发展成一个漂亮的科学分支——静力学，该理论可以合情合理地够格称作是超等的。现在该理论已被牛顿理论所包容。1600年左右由伽利略提出，并由牛顿将其发展成美丽的、内容丰富的理论，研究运动物体的动力学的根本观念，应该毫无疑问地纳入超等的范畴。把它应用于行星和月亮的运动时，具有惊人的可观察的精确性——其误差比一千万分之一还小。同一个牛顿的方案也以相当的精确性适用于地球以及外推到恒星和星系的范围。类似地，麦克斯韦理论在向内可达到原子和亚原子的粒子尺度，向外达到大约大一万亿亿亿亿倍的星系的尺度的异乎寻常的范围内准确地成立！（在此尺度的小的那一端，麦克斯韦方程必须和量子力学的规则适当地合并在一起。）它也肯定够格被称作超等的。

爱因斯坦的狭义相对论（为庞加莱所预想并被闵可夫斯基非常精巧地表述）对允许物体以接近光速运动的现象给出了令人惊叹的准确的描述。牛顿的描述最终在这种情况下开始动摇。爱因斯坦的无与伦比漂亮的和开创性的广义相对论推广了牛顿的引力动力学理论并改善了它的精确性，继承了牛顿理论处理行星和月亮运动的所有非凡的成就。此外，它还解释了各种和旧的牛顿方案不一致的观测事实。其中一个例子（参阅272页的脉冲双星的例子）指出爱因斯坦的理论能精确到大约$1/10^{14}$。两种相对论——第二种将第一种包含了——应该明确地归到超等的类中去（其数学上的优雅几乎和其准确性一样重要地作为这分类的原因）。

由不可思议漂亮的和革命性的量子力学理论所能解释的现象的范围以及它与实验符合的精度，很清楚地表明它必须归至超等的类中去。迄今尚未找到与该理论在观测上的偏差——然而在用该理论解释许多迄今令人费解的现象方面，显示出其威力远远地超过这些。化学定律、原子的稳定性、光谱线的狭窄（参阅289页）以及非常特别的花样，超导的零电阻的古怪现象以及激光的行为仅仅是其中的几个例子。

我给超等的分类立下了很高的标准，但这是我们在物理中已经习惯了的。那么，对于最近代的理论能说些什么呢？以我的观点看，恐怕其中只有一种或许够格被称为超等的，并且它还不是特别新的：即所谓的量子电动力学（或QED）。它是由约旦、海森伯和泡利提出，1926—1934年由狄拉克所表述，最终在1947—1948年由贝特、费恩曼、施温格以及朝永加以改进使之可以应用。这个理论是狄拉克将量子力学、狭义相对论、麦克斯韦方程以及制约电子自旋和运动的基本方程结合在一起的结果。总的来说，该理论缺乏早先的许多超等理论的令人信服的精巧和一致性，但它的资格在于真正惊人的准确性。特别值得一提的结果是它给出了电子磁矩的值。（电子的行为类似于一个自旋的电荷的微小磁铁。此处“磁矩”即是这小磁铁的强度。）由QED计算出的这一小磁矩的值为1.00115965246（以某一单位测量——误差大约在最后两位小数上的20），而最近的实验值为1.001159652193（误差大约在最后两位小数上的10）。正如费恩曼所指出的，其精确度等效于从纽约到洛杉矶之间相差一根头发的宽度！我们没有必要在此了解该理论。为了完整起见，我将在下一章的结尾简单地提到它的一些重要的特征[31]。

我要将一些现代理论放到有用的范畴中去。有两种理论虽然在这里不需要，却值得提及。第一个是称为强子（质子、中子、介子等组成原子核——或更准确地讲“强相互作用”的粒子）的亚原子粒子的盖尔曼——兹维格夸克模型以及描述它们之间相互作用的详细的（后期的）称为量子色动力学或QCD的理论。其思想是，所有强子都由称作“夸克”的部分组成，夸克之间以从麦克斯韦理论的某种推广（称为杨——米尔斯理论）的方式进行相互作用。第二种理论是由格拉肖、萨拉姆、瓦尔德和温伯格提出的，它又是利用杨——米尔斯理论将电磁力和描述放射性衰变的“弱”作用结合起来。该理论对所谓的轻子（电子、μ子、中微子；还有W粒子和Z粒子——所谓的“弱相互作用”的粒子）做出统一描述。这两种理论有好的实验支持。但是由于种种原因，这些理论远不如人们期望的像QED那么清爽，而且它们目前的观测精度以及预言能力离开超等类的惊人的标准还非常远。有时将这两种理论（第二种还包含QED）称作标准模型。

最后，还有另一种我相信至少可归于有用的范畴的理论。这就是称为宇宙的大爆炸起源的理论[32]。此理论在第7章和第8章的讨论中将起重要的作用。

我认为没有更多的理论属于有用的 [2] 范畴。现代（或近代）有许多盛行的观念。它们除了“GUT”理论（以及某些从它导出的观念，诸如“暴胀模型”，参阅587页的注释13）外还有：“卡鲁查——克莱因”理论、“超对称”（或“超引力”）以及还极其时髦的“弦”（或“超弦”）理论。依我之见，所有这些都毫无疑义地属于尝试类中（参阅*Barrow 1988*，*Close 1983*，*Davies and Brown 1988*，*Squires 1985* ）。在有用和尝试类之间的重大差别是后者没有任何有意义的实验支持 [3] 。但是这并不是说，其中不会有一个将戏剧性地升格为有用的甚至超等的范畴的理论。其中某些的确包含有许多相当有前途的、富有创见的思想，但是，可惜迄今仍然没有得到实验的支持，而只停留在观念阶段。尝试类是一个非常宽广的范畴。它们其中有些牵涉到包括能导致新的实质性的理解上的进步的基因，同时我认为其他的一些肯定是误导的或做作的。（我曾经受不了诱惑，试图从可尊敬的尝试类中分出称作误导的第四类——但是后来我想还是不分的好，因为我不想失去我的一半朋友！）

超等的理论主要是古代的，人们不必为此感到惊讶。在整个历史上一定有过多得多的归于尝试类的理论，但是多数都被遗忘了。与此相似，许多有用类的理论后来也被湮没了；但是也还有一些被吸收到后来归于超等类的理论中。让我们考虑一些例子。在哥白尼、开普勒和牛顿提出优越得多的方案之前，古希腊人提出过一个十分精巧的行星运动的称作托勒密系统的理论。按照这一方案，行星的运行由圆周运动的复杂组合所制约。它能相当有效地做预言，但是在需要更高的精度时，变得越来越繁复。今天我们看来，托勒密系统的人为因素显得非常突出。这是一个有用理论（实际上大约用了2000年）后来整个退出物理理论的极好例子，虽然它曾在历史上起过很重要的组织作用。相反地，开普勒的辉煌的椭圆行星运动的观念便是从有用的理论变成我们能见到的最终成功的例子。化学元素的门捷列夫周期表是另一个例子。它们并没有提供具有“惊人”特征的预言方案，但是后来成为从它们成长出来的超等的理论的“正确”的推论（分别为牛顿动力学和量子理论）。

在以后的章节中，我不再对仅仅归于有用的和尝试的范畴的现代理论多加讨论。因为超等理论已足够讨论的了。我们有这等理论，并能以非常完整的方式理解生活其中的世界，确实是非常幸运的。我们最终必须决定，甚至这些理论是否足够丰富到能制约我们头脑和精神的作用。我将依序触及这些问题；但目前让我们先考虑超等理论并深入思考它们和我们目的相关联之处。

欧几里得几何

欧几里得几何即是我们在中学当作“几何”学习的学科。然而，我预料大部分人会将其视作数学，而不视作物理。当然，它也是数学。但是，欧几里得几何绝不是仅有的可以想得出的数学几何。欧几里得传给我们的特殊几何非常精确地描述了我们生活其间的世界的物理空间，但这不是逻辑的必然——它仅仅是我们物理世界的（几乎准确的）被观察的特征。

的确还存在另外称作罗巴切夫斯基（或双曲）的几何[33]，它大部分方面非常像欧几里得几何，但还具有一些有趣的差别。例如，我们记得在欧几里得几何中任意三角形的三个角的和为180°。在罗巴切夫斯基几何中，这个和总是比180°小，并且这个差别和三角形的面积成比例（图5.1）。

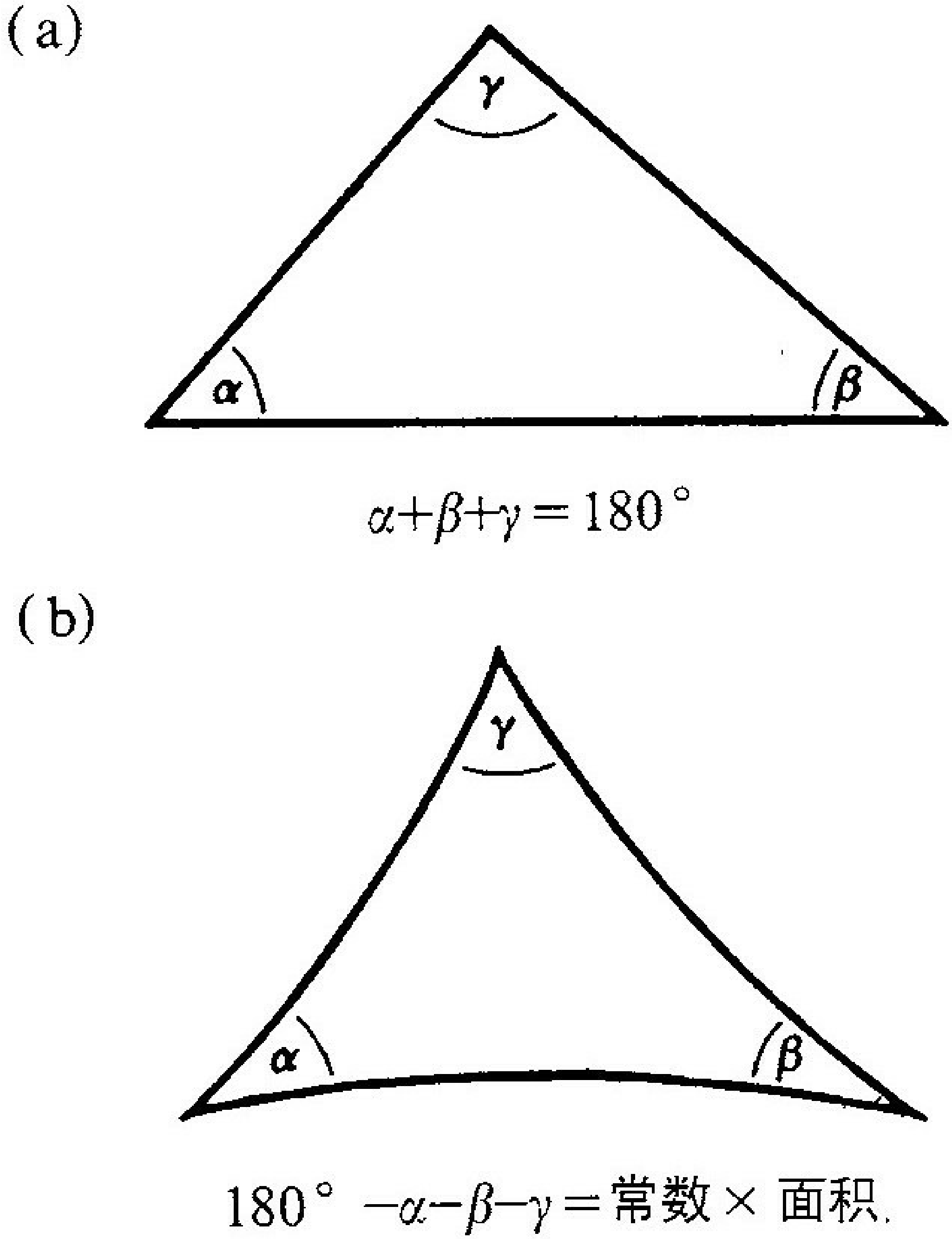


图5.1（a）欧几里得空间中的一个三角形（b）罗巴切夫斯基空间中的一个三角形

著名的荷兰艺术家毛里兹·C.伊歇为这种几何给出了一种非常精细和准确的表象（图5.2）。按照罗巴切夫斯基几何，所有的黑鱼具有相同的大小和形状；类似地，白鱼亦是如此。不能将这种几何在通常的欧几里得平面上完全精密地表达出来，所以在圆周边界的内缘显得非常拥挤。想象你自身位于该模型的某一靠近边界的地方，罗巴切夫斯基几何

使你觉得就像位于中间或任何其他地方一样。按照这一欧几里得表象，该模型的“边界”正是罗巴切夫斯基几何中的“无穷远”。此处边界圆周根本不应该被看成罗巴切夫斯基空间的一部分——在圆周之外的任何其他的欧几里得区域就更不是了。（这一罗巴切夫斯基平面的天才表象应归功于庞加莱。它突出的优点在于，非常小的形状在此表象中不被畸变——只不过它的尺度被改变。）该几何中的直线（伊歇鱼就是沿着其中某些直线画出的）即为与边界圆周作直角相交的圆弧。

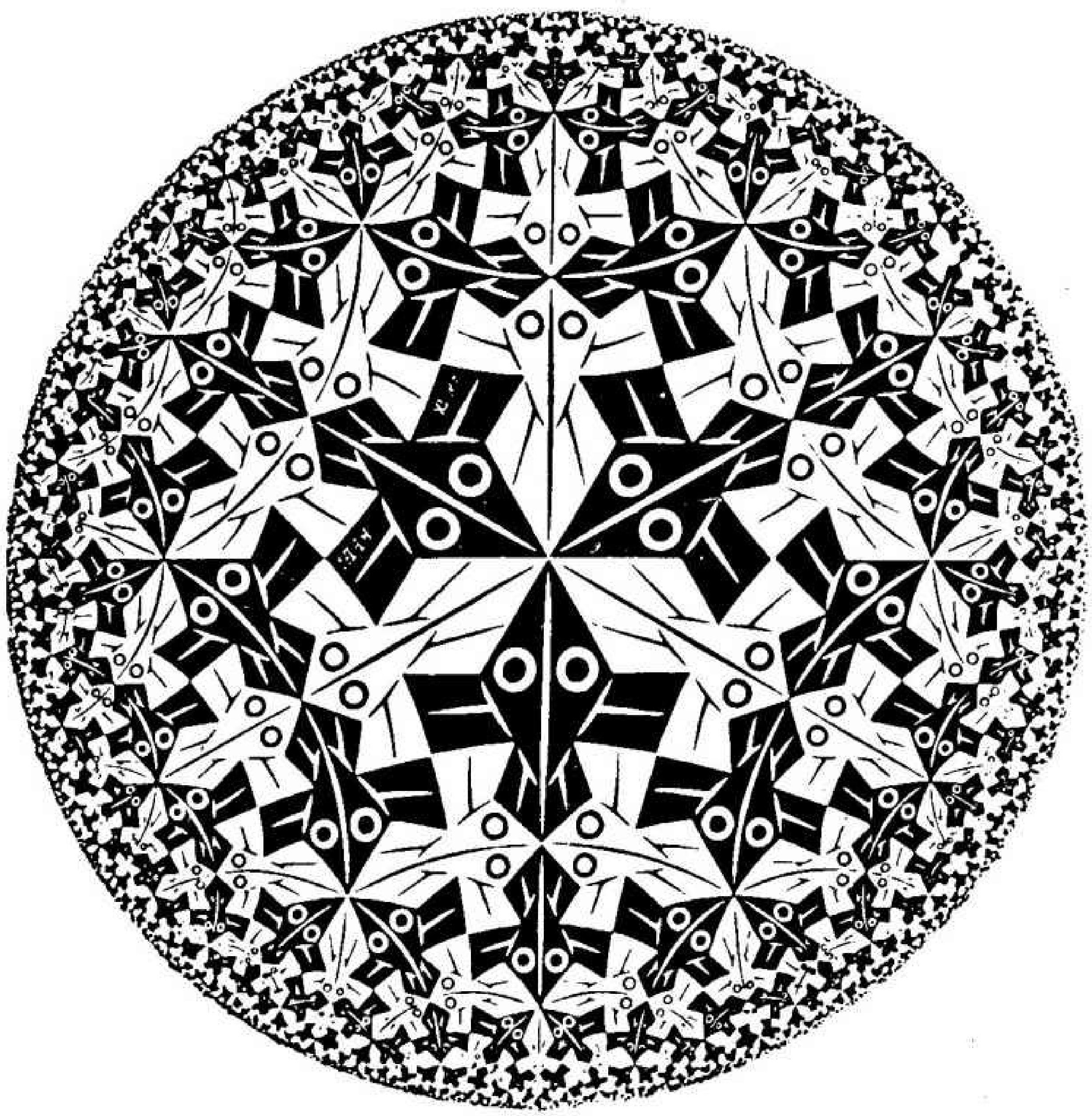


图5.2　罗巴切夫斯基空间的伊歇图（所有黑鱼和白鱼都认为是全等的）。[]1958 M.C.Escher/Cordon Art-Baarn-Holland

我们世界在宇宙学的尺度下，实际上很可能是罗巴切夫斯基空间（参阅第7章411页）。然而，在这种情形下，三角形亏角和它的面积的比例系数必须是极为微小。在通常的尺度下，欧几里得几何是这种几何的极好的近似。事实上正如我们在本章将要看到的，爱因斯坦的广义相对论告诉我们，在比宇宙学尺度小相当多的情形下，我们世界的几何的确与欧几里得几何有偏差（虽然是以一种比罗巴切夫斯基几何更复杂的“更无规”的方式），尽管这偏差在我们直接经验的尺度下仍是极为微小的。

欧几里得几何似乎精确地反映了我们世界“空间”的结构的这一事实，作弄了我们（以及我们的祖先），使我们以为几何是逻辑所必需的，或以为我们有种先天的直觉的领悟，欧几里得几何必须适用于我们在其中生活的世界。（甚至伟大的哲学家伊曼努尔·康德也作此断言。）只有爱因斯坦在许多年以后提出的广义相对论真正地突破了欧几里得几何，欧几里得几何远非逻辑所必需的，它只是该几何如此精确地（虽然不是完全准确地）适合于我们物理空间结构的经验的观测事实！欧几里得几何确实是一个超等的物理理论。这是它作为纯粹数学的一部分的精巧性和逻辑性以外的又一个品质。

在某种意义上，这和柏拉图（约公元前360年，大约在欧几里得著名的《原本》一书出版之前50年左右）信奉的哲学观点相差不远。依柏拉图观点，纯粹几何的对象——直线、圆周、三角形、平面等——在实际的物理世界中只能近似地得到实现。而那些纯粹几何在数学上的精确对象居住在一个不同的世界里——数学观念的柏拉图的理想世界中。柏拉图的世界不包括有可感觉的对象，而只包括“数学的东西”。我们不是通过物理的方法，而是通过智慧来和这个世界接触。只要人的头脑沉思于数学真理，用数学推理和直觉去理解，则就和柏拉图世界有了接触。这个理想世界被认为和我们外部经验的物质世界不同，虽然比它更完美，但却是一样地实在。（回顾一下我们在第3章127页和第4章147页关于数学概念的柏拉图实在性的讨论。）这样，可以单纯地用思维来研究欧几里得几何，并由此推导其许多性质，而外部经验的“不完美的”世界不必要刚好符合这些观念。基于当时十分稀少的证据，柏拉图以某种不可思议的洞察力预见到：一方面，必须为数学而研究数学，不能要求它完全精确地适用于物理经验的对象；另一方面，实际的外部世界的运行

只有按照精确的数学——亦即“智慧接触得到的”柏拉图理想世界才能最终被理解。

柏拉图在雅典创建了学园以推动这种观念。极富影响的著名的哲学家亚里士多德即为其中之出类拔萃者。但是我们要在这里论及另一位比亚里士多德名望稍低的学园成员，即数学家兼天文学家欧多克索斯。依我看来，他是一位更优秀的科学家，也是古代最伟大的思想家之一。

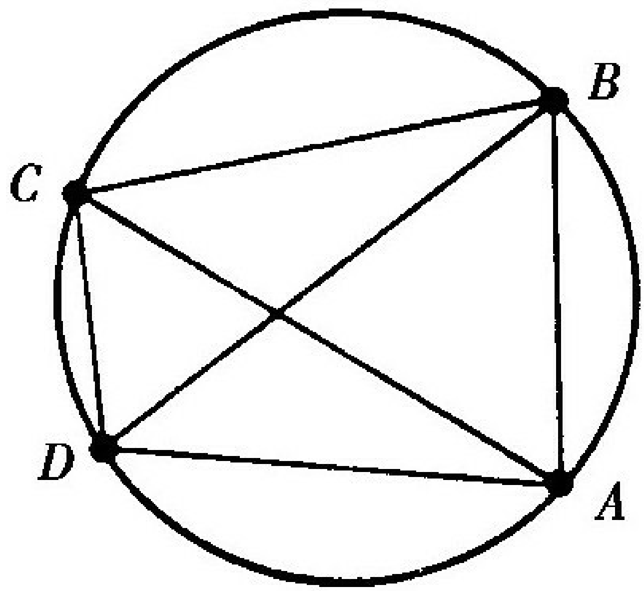
欧几里得几何中有一基本的、微妙的，并的的确确最重要的部分，那就是实数的引进，虽然今天我们几乎不认为它是几何的（数学家宁愿将它称作“分析”的，而非“几何”的）。因为欧几里得几何研究长度和角度，所以必须了解用何种“数”来描写长度和角度。新观念的核心是在公元前4世纪由欧多克索斯（约公元前408至前355年）[34]提出的。由于毕达哥拉斯学派发现了像$\sqrt{2}$这样的数不能被表达成分数，使得希腊几何陷入了“危机”之中（参阅第3章第104页）。将正方形的对角线，以其边长来度量时就必然出现$\sqrt{2}$这个数。对于希腊人来说，为了用算术的定律来研究几何量，将几何测量（比）按照整数（比）来表示是很重要的。欧多克索斯的基本思想是提供一种以整数表达线段之比的办法（也就是实数）。他依赖于整数的运算提出了决定一个线段之比是否超过另一个线段之比，或两者是否完全相等的判据。

该思想可概述如下：如果a, b，c和d是4条线段，则断定比例a/b大于比例c/d的判据是：存在整数M和N，使得a增大到N倍超过b增大到M倍，而同时d增大到M倍超过c增大到N倍[35]。可用相应的判据来断定a/b是否比c/d小。所寻求的使a/b和c/d相等的判据也就是前两个判据都不能满足！

直到19世纪，戴德金和魏尔斯特拉斯等数学家才发展出完全精确的抽象的实数数学理论。但是他们的步骤和欧多克索斯早在22个世纪以前已经发现的思路非常相似！我们在此没有必要描述这个现代发展。在第3章第107页我已给出了这个理论的模糊暗示。但是，为了更容易表达，我宁愿在这里用更熟悉的小数展开的方法来讨论实数理论。（这种展开实际是在1585年才由斯蒂文引进的。）必须指出，虽然我们很熟悉小数表达方式，但希腊人却对此无知。

然而，在欧多克索斯设想和戴德金——魏尔斯特拉斯设想之间有一

个重大差别。古希腊人认为实数是由几何量（之比）产生的东西，当作“实在”空间的性质。希腊人用算术来描述几何量是为了要严格地处理这些量以及它们的和与积——亦即古人这么多美妙几何定理的要素的先决条件。（我在图5.3画出并解释了杰出的托勒密定理——虽然托勒密比欧多克索斯要晚许久才发现它——该定理和一个圆周上的4点之间的距离相关，它很清楚地表明了和与积都是需要的。）历史证明欧多克索斯判据极其富有成果，尤其是它使希腊人能严格地计算面积和体积。



$$AB \cdot CD + AD \cdot BC = AC \cdot BD$$

图5.3　托勒密定理

然而，对于19世纪尤其是当代的数学家而言，几何的作用已被改变了。古希腊人，尤其是欧多克索斯，认为“实”数是从物理空间的几何中抽取出来的东西。现在我们宁愿认为在逻辑上实数比几何更基本。这样的做法还可以允许我们建立所有不同种类的几何，每一种几何都是从数的概念出发。（其关键的思想是16世纪由费马和笛卡儿引进的坐标几何。坐标可用来定义其他种类的几何。）任何这种“几何”必须是逻辑相容的，但不必和我们经验的物理空间有任何直接的关联。我们似乎感知的特别物理几何是经验的理想化（例如，依赖于我们将其向无穷大或无

穷小尺度的外推，参阅第3章第112页）。但是现代的实验已足够精密，以至于我们必须接受“经验的”几何的确和欧几里得观念有差别的这一事实（参阅272页）。这种经验和从爱因斯坦广义相对论推导的结果相一致。然而，尽管我们的物理世界的几何观点起了变化，欧多克索斯23世纪之久的实数概念在实质上并没有改变。它对爱因斯坦理论正如对欧几里得理论一样重要。其实，迄今为止它仍然是一切严肃物理理论的重要部分。

欧几里得《原本》的第五部基本上是关于欧多克索斯“比例论”的阐述。这对整本书而言是极为重要的。全书首版于公元前300年的《原本》的确必须列为有史以来最具深远影响的著作之一。它成为后来的几乎所有科学和数学思想的舞台。它全部是由一些被认为空间的“自明”性质，亦即清楚叙述的公理出发演绎而来，其中许多重要推论根本不是显而易见的，而是令人惊异的。无疑地，欧几里得的著作对后世科学思想的发展具有深刻的意义。

阿基米德（公元前287—前212）无疑是古代最伟大的数学家。他天才地利用欧多克索斯的比例论，计算出诸如球体，或者更复杂的牵涉到抛物线和螺线的许多不同形体的面积和体积。今天我们可以用微积分十分容易地做到这些。但是我们要知道，这是比牛顿和莱布尼茨最终发现微积分早19个世纪的事！（人们可以说，阿基米德已经通晓微积分的那一多半——亦即积分的那一半！）阿基米德的论证，甚至以现代的标准看，也是毫无瑕疵的。他的写作深深地影响许多后代的数学家和科学家，最明显的是伽利略和牛顿。阿基米德还提出了静力学的（超等的？）物理理论（亦即制约平衡的物体，诸如杠杆和浮体的定律）。他用类似于欧几里得发展几何空间和刚体几何的科学方法，将其发展成演绎的科学。

阿波罗尼奥斯（约公元前262—前200）是我必须提及的一位阿基米德的同时代人。他是一位具有深刻洞察力的、伟大的、天才的几何学家。他关于圆锥截线（椭圆、抛物线和双曲线）的研究极大地影响了开普勒和牛顿。令人惊异的是，这些截线的形状刚好是描述行星轨道所必需的！

伽利略——牛顿动力学

对运动的理解是17世纪科学的根本突破。古希腊人对静态的物理——刚性的几何形状或处于平衡的物体（此时所有的力都平衡，因而没有运动）理解得很透彻。但是他们对制约实际运动的物体的定律并没有很好的概念。他们所缺少的是一个好的动力学理论，亦即自然实际上控制物体的位置从第一时刻到下一时刻变化的完美方式的理论。其部分原因（绝非全部）则是没有测量时间的足够精密的手段，亦即没有相当好的“钟表”。为了给位置变化定时以及确定物体的速度和加速度，人们必须有钟表。因此，1583年伽利略观察到摆能作为计时的可靠手段的这个事件对他（甚至对整个科学！）极具重要性，因为这样一来运动的计时就变准确了 [4] 。随着55年后的1638年伽利略《对话》一书的出版诞生了新的学科——动力学——开始了从古代神秘主义到现代科学的转化！

让我仅仅列举伽利略提出的4个最重要的物理观念，第一是作用在物体上的力决定的是它的加速度，而不是速度。此处“加速度”和“速度”的含义是什么呢？粒子——或物体上的某点——的速度是该点位置相对于时间的变化率，速度通常是一个矢量，亦即必须同时考虑其方向和大小的量（否则我们用“速率”这一术语，见图5.4）。加速度（又是一个矢量）是速度相对于时间的变化率——这样加速度实际上是位置相对于时间的变化率的变化率！（这对于古人来说实在太难为了！他们既缺乏可胜任的“钟表”，又不具备与变化率相关的数学概念。）伽利略断言，作用在物体的力（在他的情形下是指重力）制约物体的加速度，而非直接制约其速度——正和古代人（例如亚里士多德）所相信的不一样。

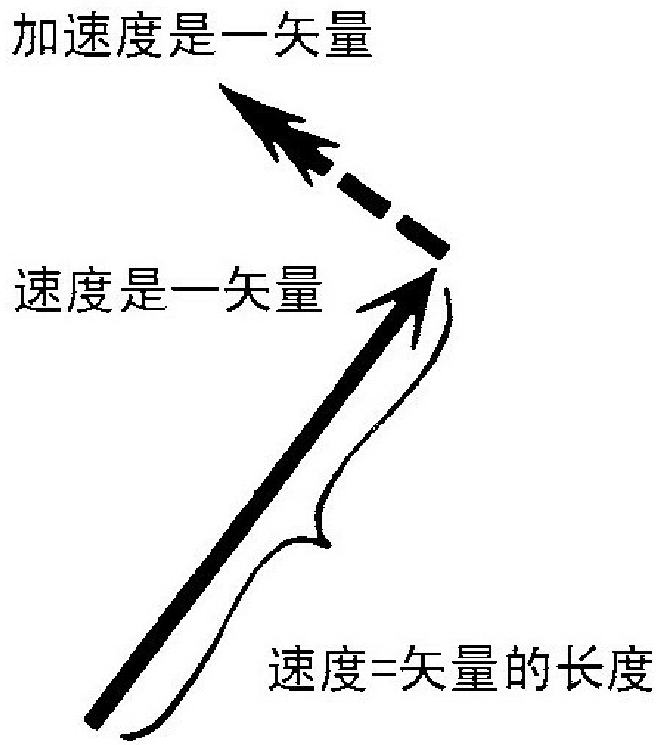


图5.4　速度、速率和加速度

特别是当不存在外力时，速度必须是常数——因此，在直线上作的恒常运动应是没有外力作用的结果（牛顿第一定律）。自由运动着的物体继续其匀速运动，而不必施加外力去维持它。伽利略和牛顿发展的动力学定律的一个推论是，直线匀速运动和静止状态亦即不运动在物理上完全不可区分：不存在一种局部的方法，将匀速运动从静止中区别开来！伽利略关于这点特别清楚（甚至比牛顿还清楚）。他以海上的航船作例子对此作了非常形象的描绘（参阅Drake 1953，P182—P183）：

把你和某位朋友关在某艘大船的甲板下的主舱里，和你一道的还有一些苍蝇、蝴蝶和其他飞行的小动物。一些鱼在一大碗水中自由自在地游着；水一滴一滴从悬挂着的瓶子落到下面的一个大器皿里。当船静止时，仔细观察这些小动物如何以同样的速率向船舱的所有方向飞行。鱼儿不辨方向地游着，水滴落到下面的器皿中；……在仔细地观察了这一切以后……让船以你想要的速度行驶。只要其运行是均匀的，并且不让它作这样那样的摇动，你就会发现，不但所有提及的现象没有丝毫变化，而且你根本就不知道船是在行驶，还是在静止不动……正如早先那

样，小水滴落到下面的器皿中去，而不向船尾的方向飘去，虽然就在水滴在空气中的时间间隔里，船已经向前走了船身长度好几倍的距离。水中的鱼向前游动并不比向后更费力，同样轻松地向放在碗的任何方向的边缘上的鱼饵游去。还有，蝴蝶和苍蝇毫无异样地继续飞向四方。似乎它们为了避免落后，在空中随着船作长途旅行后感到疲劳，最后聚集到船尾的现象从未发生过！

这个被称为伽利略相对论原理的惊人事实，在使哥白尼观点具有动力学意义上十分关键。尼古拉·哥白尼（1473—1543）[以及古希腊天文学家亚里斯塔哥斯（约公元前310—前230）——不要和亚里士多德相混淆！—比他早18个世纪]提出了日心说，即太阳处于静止状态，而地球在沿自己的轴自转的同时绕着太阳公转，公转速度为每小时10万千米。为何我们没有感觉到这种运动？在伽利略提出动力学理论之前，这的确是哥白尼观点的深深的困惑。如果更早先的“亚里士多德式”的动力学观点是正确的话，即在空间中运动的系统的实在速度要影响其动力学行为，那么地球的运动对我们就会有直接明显的效应。伽利略相对论弄清了，何以地球在运动，而同时我们却不能直接感觉到它的原因[36]。

值得指出的是，在伽利略相对论中，“静止”的概念并无任何局部上的意义。它对人类的空间——时间观念已经具有显著的含义。我们直观的空间——时间图像是，“空间”构成了物理事件在其中发生的舞台。物理对象在某一时刻可处于空间的某一点，在后一时刻可处于同一个，或另一个不同的空间的点。我们想象空间中的点可以从一个时刻维持到另一个时刻。这样，一个物体实际上是否改变其空间位置的说法就具有意义。但是，伽利略相对论指出，不存在“静止状态”的绝对意义；所以，“在不同时间的空间的同一点”的说法是毫无意义的。某一时刻的物理经验的欧几里得三维空间的哪一点是我们的欧几里得三维空间另一时刻的“同一点”呢？没有办法找到。对应于每一时刻我们似乎必须有一个完全“新”的欧几里得空间！考虑具有物理实在性的四维空间——时间图就会使这一层意思明了（图5.5）。不同时刻的欧几里得三维空间的确被分开，但所有这些空间合并在一起构成了完整的四维的空间——时间图。在空间——时间中进行匀速直线运动的粒子的历史是一条直线（称为世界线）。以后在讨论爱因斯坦相对论时我还会回到空间——时间以及运动的相对性的问题上来。我们将发现在那种情形下对四维维数的论证会更加有力。

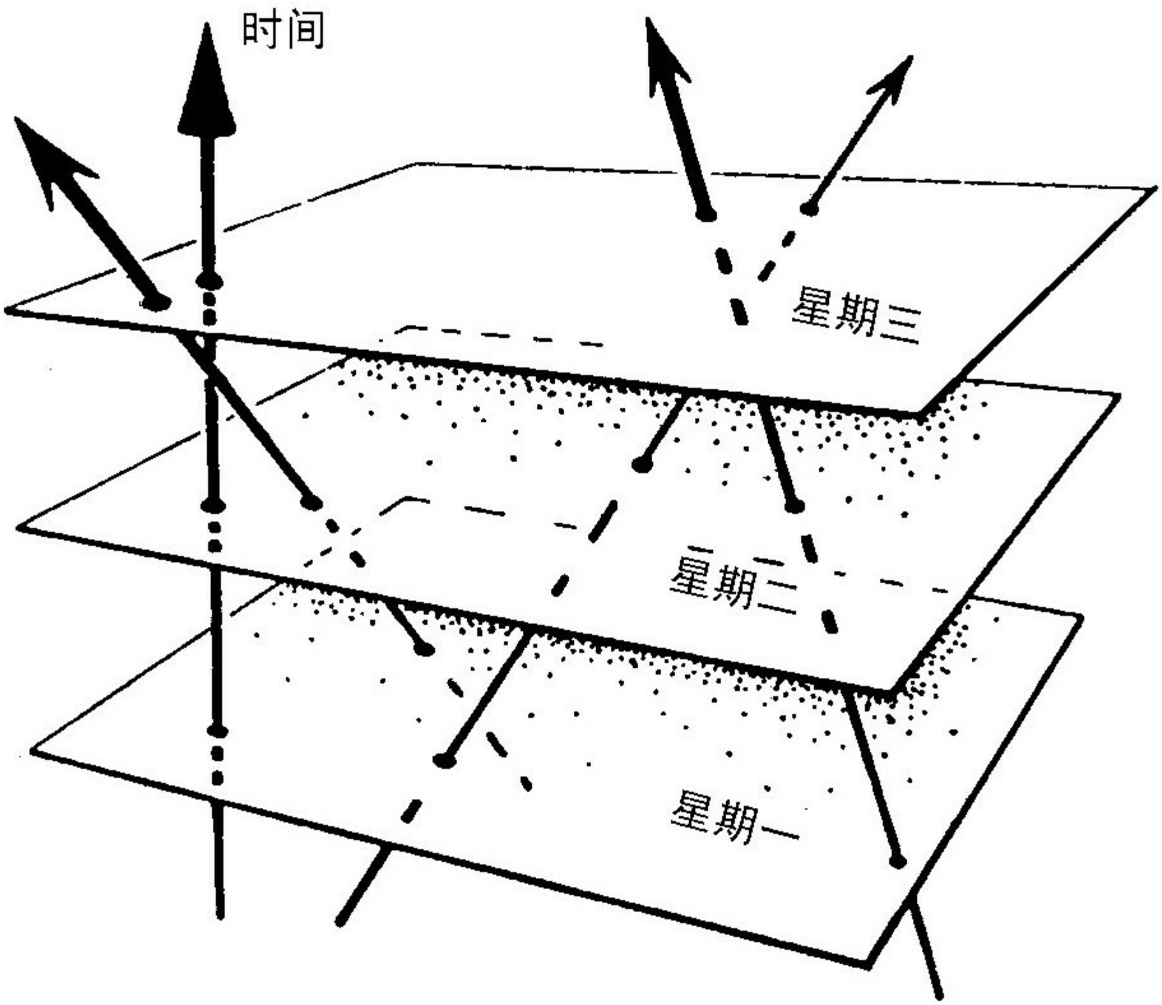


图5.5　伽利略空间——时间：匀速运动的粒子可用直线标出

伽利略的第三个伟大洞察是开始理解能量守恒。伽利略主要关心物体在重力下的运动。他注意到，如果从一静止状态释放一个物体，则不管它是简单地落下，还是随一个任意长度的摆振动，或是沿着一个光滑斜面滑下，其速率只依赖于它从释放之处下落的垂直距离。正如我们现在所说的，储存于超过地面的高度的能量（引力势能）会转换成它的运动的能（只依赖于物体速率的动能）。反之亦然，但总能量既不损失也不增加。

能量守恒定律是一个非常重要的物理原则。它不是物理学的一个独立要求，而是我们很快就要讨论的牛顿动力学定律的推论。笛卡儿、惠更斯、莱布尼茨、欧拉以及开尔文等人几个世纪来的努力，使这一定律的表述越发清晰。在本章的后面部分以及第7章，我们将要再回到这个

问题上来。如果把能量守恒定律和伽利略的相对论原理相结合，我们就能得到更多的相当重要的守恒定律：质量和动量守恒。粒子的动量是它的质量和速度的乘积。火箭的推进即是动量守恒的众所周知的例子之一，火箭往前动量的增加恰好和（更轻的，但是更急速的）废气往后的动量相平衡。枪的后坐力也是动量守恒的一个表现形式。牛顿定律的进一步推论是角动量守恒，角动量守恒是描写一个系统的自旋的不变性，地球绕自己的轴自旋以及网球的自旋都是依靠它们的角动量守恒来维持的。组成任何物体的每一个粒子都对该物体的总角动量有贡献，这贡献等于它的动量与它离开中心的垂直距离的乘积。（自旋转物体只要变紧凑，其角速度就会增加，即是其中的一个推论。滑冰者和马戏团高架秋千艺术家经常表演的令人惊叹而熟悉的动作也起源于此。他们经常利用收回手臂或腿的动作使旋转速度自动增加。）在后面的内容中我们将会看到质量、能量、动量以及角动量都是重要的概念。

最后，我应该让读者回顾一下伽利略的先知的洞察力，那就是当大气摩擦力不存在的时候，在重力作用下所有物体都以同一速率下落。（读者也许会回想起他从比萨斜塔上同时释放不同物体的著名故事。）3个世纪以后，正是这一个洞察导致爱因斯坦将其相对论原理推广到加速参考系统，从而为他的非凡的引力的广义相对论提供了基石，这在本章的结尾处将会看到。

在伽利略创立的令人印象深刻的基础上，牛顿建立了绝顶庄严华美的大教堂。牛顿指出了物体行为的定律。第一和第二定律基本上是伽利略给出的：如果没有外力作用到一个物体上，则物体将继续其匀速直线运动；如果有外力作用到上面，则物体的质量乘以它的加速度（亦即其动量变化率）等于这个力。牛顿本人的一个特殊的洞察，在于意识到还需要第三定律：物体A作用在物体B上的力，刚好和物体B作用到物体A上的力大小一样而方向相反（“每一个作用必有其大小一样方向相向的反作用”）。这就提供了基本的框架。“牛顿宇宙”是由在服从欧几里得几何定律的空间中运动的粒子所组成。作用到这些粒子上的力决定了它们的加速度。每一个粒子所受的力是由所有其他粒子分别贡献到该粒子的力利用矢量加法定律相加而得到的（图5.6）。为了很好地定义这个系统需要一些规则，这些规则可以告诉我们从另一个粒子B作用到粒子A的力是什么样子的。通常我们需要该力沿着AB之间的连线作用（图5.7）。如果该力是引力，则A和B之间的力是互相吸引的，其强度和它们质量乘积成正比，而和它们之间的距离的平方成反比：亦即平方反比

律。对于其他种类的力，其依赖于位置的方式可与此不同，也可能决定于粒子质量以外的其他性质。

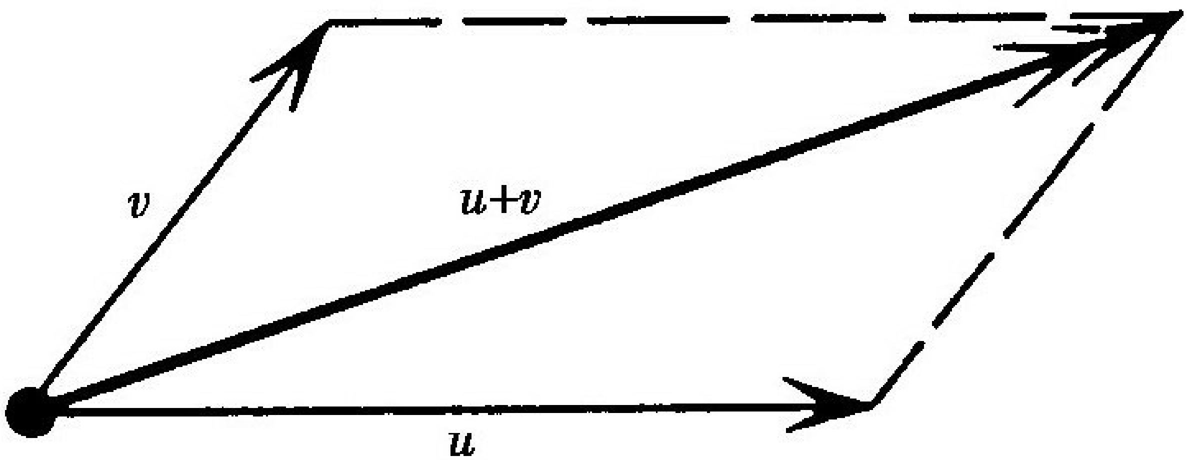


图5.6　矢量加法的平行四边形定律

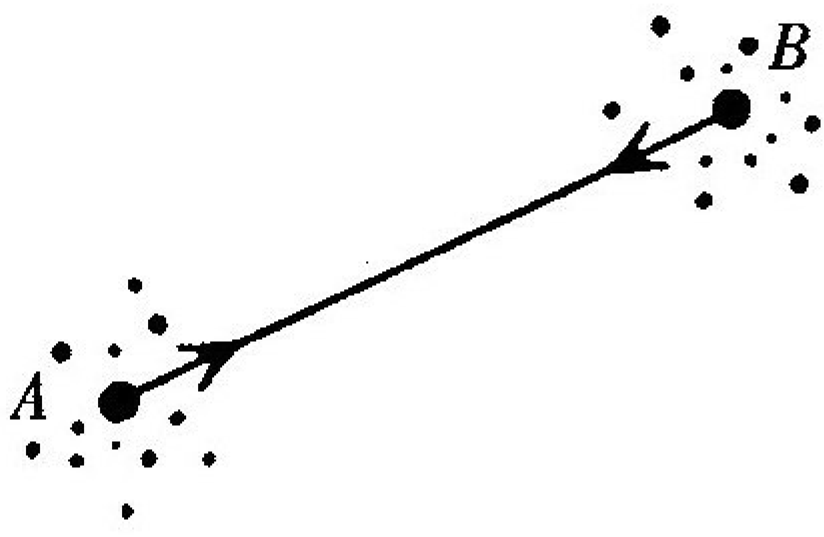


图5.7　两粒子之间的力是沿着它们之间连线的方向（由牛顿第三定律，B作用到A的力总是和A作用到B的力大小相等且方向相反）

伽利略的一位同时代人，伟大的约翰斯·开普勒（1571—1630）注意到，行星绕太阳公转的轨道是椭圆而不是圆周（太阳总是处于该椭圆的一个焦点上，而不在其中心）。他还给出了制约行星作此椭圆运动的

速率的其他两个定律。牛顿能够从他自己的一般理论（以及引力的平方反比律）推导出开普勒三定律。不仅如此，他还对开普勒的椭圆轨道做了各种细节上的修正，诸如春秋分日点的进动（许多世纪以前的希腊人已注意到这些地球旋转轴方向的这种极慢的运动）。为了取得所有这些成就，牛顿就必须发展除微积分之外的许多数学手段。他惊人的成就得大大归功于其超等的数学技巧及其同等超人的物理洞察力。

牛顿动力学的机械论世界

如果已知特定的力的定律（例如引力的平方反比律），则牛顿理论就表达成一组精密的确定的动力学方程。如果各个粒子在某一个时刻的位置、速度和质量是给定的，则它们随后任何时间的位置、速度（以及质量——这被当作常数）就在数学上确定，这种牛顿力学的世界所满足的决定论形式对哲学思维产生了（并正在产生着）深远的影响。让我们更仔细地考察牛顿的决定论。它对“自由意志”有何含义呢？一个严格的牛顿世界能包含精神吗？甚至牛顿世界能包含计算机器吗？

让我们先明确一下什么是世界的“牛顿”模型。例如，我们可以认为组成物体的所有粒子是数学的，亦即没有尺度的点。另外的办法是将它们当作球状的刚性球。无论如何，我们都必须假定知道力的定律，例如，牛顿引力论中的引力的平方反比律。我们还要对自然的其他力，比如电力和磁力（威廉·吉尔伯特在1600年首先仔细研究过）以及现代已知将粒子（质子和中子）绑在一起形成原子核的强核力的定律也表述出来。电力正和引力一样满足平方反比律，但类似的粒子互相排斥（而不像引力那样互相吸引）。这里不是粒子的质量，而是它们的电荷决定它们之间电力的强度。磁力和电力一样也是“平方反比的”[37]，但是核力以相当不同的形式随距离而变化。在原子核中当粒子相互靠得紧密时核力极大，而在更大距离下则可以忽略不计。

假定我们采用刚体圆球的模型，并要求两个球碰撞时，它们即完全弹性地反弹。也就是说，它们如同两个完好的撞球那样，在能量（或总动量）没有损失的情况下分离开。我们还必须明确指明两球之间的作用力。为了简单起见，我们可以假定任两球之间的作用力都沿着它们中心的连线，其大小为该连线长度的给定的函数。（由于牛顿的一个出色的定理，此假设对牛顿引力自动成立。对其他力的定律，这可当成一个协

调的要求而加上的条件。）如果刚体只进行成对碰撞，而不发生3个刚体或更多个刚体的碰撞，则一切都定义得很好，而且结果会连续地依赖于初始条件（亦即只要初态的变动足够小，则能保证结果变化也很小）。斜飞碰撞的行为是两球刚好相互错过的行为的连续过渡。但在三球或多球碰撞的情形下就产生了新问题。例如，如果三球A、B、C一下子跑到一块，那A、B先碰撞，紧接着C和B碰撞，或A、C先碰撞，紧接着B和A碰撞，情况就很不一样（图5.8）。在我们的模型中，只要有三体碰撞发生就存在非决定性！只要我们愿意，就可以用“极不可能”的理由简单地将三体碰撞或多体碰撞的情形排除掉。这就提供了一种相当一致的方案，但三体碰撞的潜在问题表明终态将以不连续的方式依赖于初态。

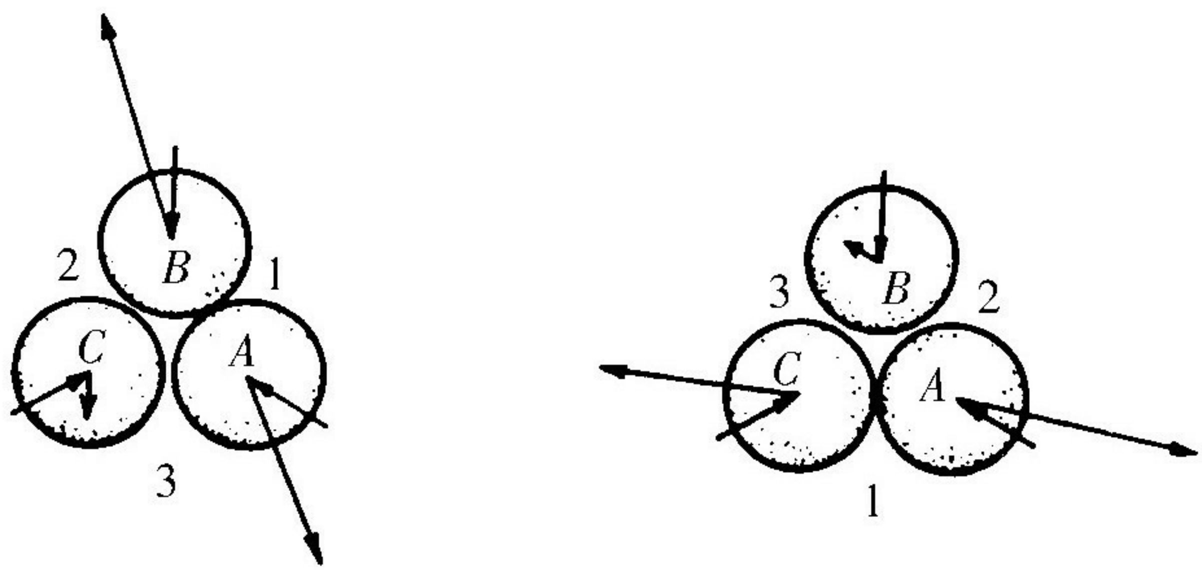


图5.8　三体碰撞。最后的行为关键地决定于哪两球先碰撞，这样使得结果不连续地依赖于起因

这有点使人不满意，我们也许会更喜欢点粒子的图像。但是，为了避免某些点粒子模型引起的理论困难（当两个粒子撞到一起时出现的无限大力和无限大能量），人们必须做其他假设，诸如在短距离时粒子的相互作用力变成非常强的排斥力，等等。在这种情形下，我们可以保证任何一对粒子实际上都不会碰撞到一起。（这也使我们避免了它们碰撞时的点粒子如何运动的问题！）然而，为了直观起见，我宁愿完全按照钢球模型来讨论。看来这种“台球”图像正是大多数人下意识的实体的模型。

牛顿[5] 台球的实体模型（不管多碰撞问题）确实是一个决定论模型。此处“决定论”的含义是：所有球（为了避免某些麻烦，假定为有限个）在将来（或过去）的物理行为数学地被某一时刻的位置和速度所完全决定。这样看来，在这个台球的世界上根本没有余地让“精神”用“自由意志”的行动去影响物体的行为。我们如果还信仰“自由意志”的话，就要被迫对实际世界的如此构成方式提出质疑。

这个令人烦恼的“自由意志”问题一直徘徊在这整部书的背景里——虽然在多数情况下，我必须说只在背景里。在本章后面有一个很小却很奇特的地方牵涉到它（关于相对论中超光速信号传递的问题）。我将在第10章直接着手自由意志的问题。读者一定会对我的结果深感失望。我的确相信，这里存在一个真正的，而非想象的问题。但它是非常根本的，并且要把它表述清晰非常困难。物理理论中的决定论是一个非常重要的问题，但是我相信这只是问题的一部分。例如，这个世界很可能是决定性的，但同时却是不可计算的。这样，未来也许以一种在原则上不能计算的方式被现在所决定。我将在第10章论证，我们具有意识的头脑的行为的确是非算法的（亦即不可计算的）。相应地，我们自信所具备的自由意志就必然和制约我们在其中生活的世界的定律中某些不可计算部分紧密地纠缠在一起。是否接受这样的关于自由意志的观点，亦即给定的物理（例如牛顿）理论，是否的确是可计算的，而不仅仅是否是决定性的，是一个有趣的问题。可计算性不同于决定性——这正是我试图在本书中所要强调的。

台球世界中的生活是可计算的吗

我现在使用一个决定性的，但不可计算的“玩具宇宙模型”，来解释可计算性和决定性是不同的。我承认这是一个人为的特别例子。宇宙任何“时刻”的“态”可用一对自然数（m，n）来表示。用T_u 表示一台固定的通用图灵机，譬如在第2章（73页）定义的那一台。为了决定下一“时刻”宇宙的态，我们必须知道T_u 在m上的作用最终停止或不停止（亦即用第2章76页的记号，T_u （m）≠□还是T_u （m）=□成立）。如果它停止，则下一时刻的态为（m+1，n）。如果它不停止，则为（n+1，m）。从第2章我们知道，不存在图灵机停机问题的算法。这样就不存在去预言这个模型宇宙“将来”的算法，尽管它是完全决定性的[6]！

当然，这不能认为是一个严肃的模型。但它表明存在一个要回答的问题。我们可对任何决定性的物理理论考察其可计算性。那么，牛顿的台球世界究竟是否可计算的呢？

物理可计算性的问题部分地依赖于我们打算对此系统问哪一种问题。在牛顿台球模型中，我能想到一些可以问的问题。我对这些问题的猜测是，要弄清其答案不是一个可计算（亦即算法的）事体。球A和球B究竟会碰撞否便是这样的一个问题。其思路是，在某一特定时刻（t=0）所有球的位置和速度作为初始数据给定后，我们要知道，A和B是否会在将来的任一时刻（t＞0）碰撞？为使这个问题更明确（虽然不是特别现实），我们可以假定，所有球的半径和质量都一样，并且每一对球之间的作用力是平方反比律的。我之所以猜想这是非算法可解的问题的一个原因是，该模型有点像爱德华·弗列得钦和托马索·托弗里在1982年提出的“计算的台球模型”。在他们的模型中，球被若干堵“墙”所限制（而不是平方反比律的力）；但是它们互相以类似于我刚描述过的牛顿球那样弹性反弹（图5.9）。在弗列得钦——托弗里模型中，所有电脑的基本逻辑运算都可由球来实现。我们可以模拟图灵机的任何计算：对图灵机T_u 的特别选取规定了弗列得钦——托弗里机器的“墙”等的搭配；运动的球的初态可认为是输入磁带的信息的码，将球的终态解码就得到图灵机输出磁带的信息。这样一来，人们会特别关心这样一个问题：如此这般的图灵机会有停止之时吗？“停机”意味着球A最终和球B碰撞。我们已知这个问题不能用算法回答（78页），这事实至少暗示我前面提出的“球A最终和球B碰撞吗？”的牛顿问题也不能用算法回答。

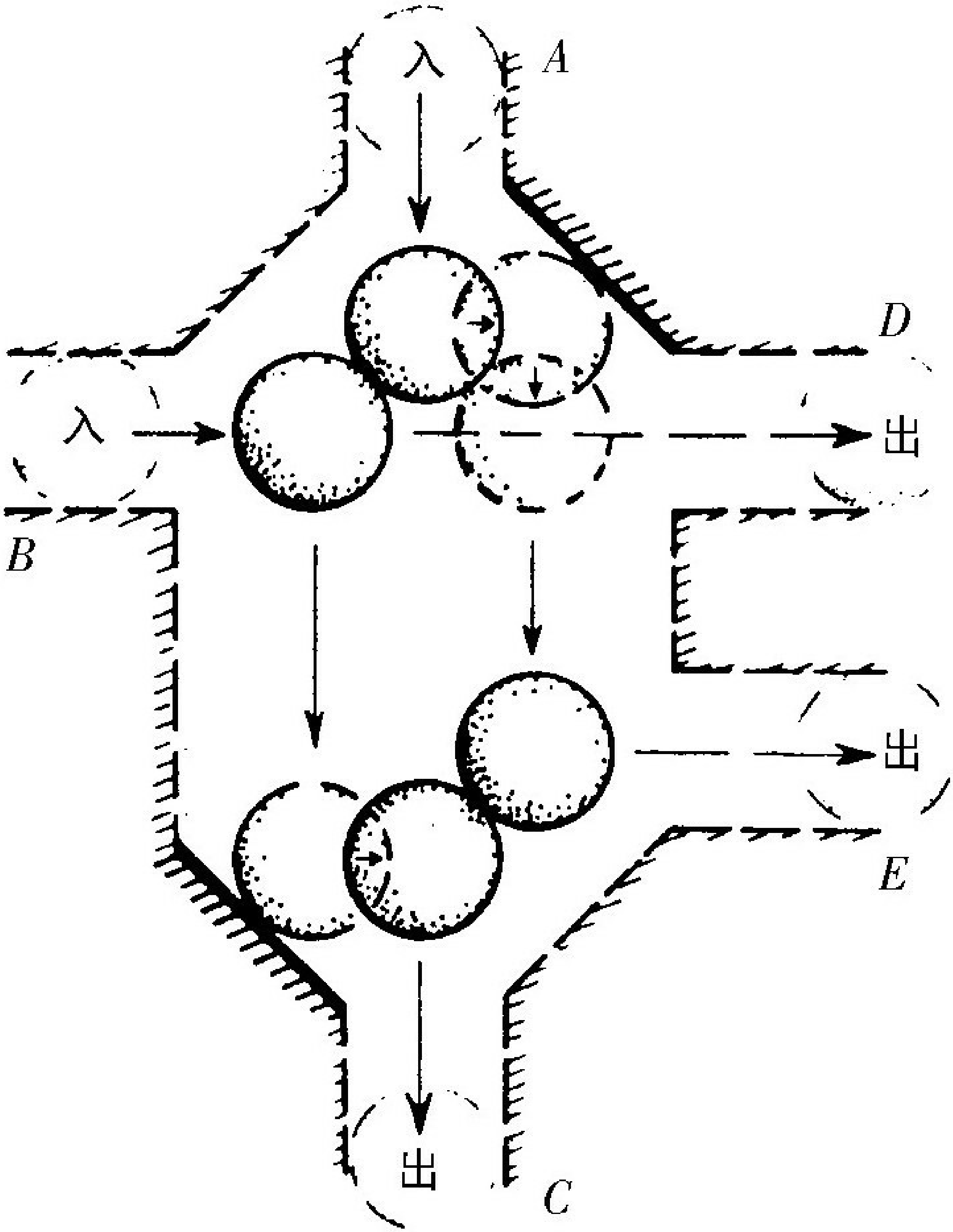


图5.9　弗列得钦——托弗里台球电脑中的一个“开关”（由A.雷斯勒提出的）。如果一个球进入B，则是否有一个球接着从D或E出来，得看是否另一球进入A中（假定A和B的同时进入）

事实上，牛顿问题比弗列得钦和托弗里提出的问题要棘手得多，后

者可依照离散参数（亦即按照诸如“球或者在通道上或者不在”的“在或不在”的陈述）来指明其状态。但在完整的牛顿问题中必须以无限的精度，按照实数的坐标而不是以离散的方式指明球的初始位置和速度。这样，我们又面临着在第4章处理关于芒德布罗集是否可递归的问题时所必须考虑的所有麻烦。当允许输入和输出数据为连续变化的变数时“可计算性”的含义是什么呢[7]？我们可以暂时假定所有初始位置和速度坐标均为有理数（虽然不能预料在t的时刻以后的有理数仍保持为有理数），而使此问题变得稍为缓和。我们知道有理数为两整数的比，所以为一可数集。我们可用有理数来任意地逼近所选择用来考察的任何初始数据了。对于有理数的初始数据，也许不存在决定A和B是否最终会碰撞的算法的猜测决不是毫无道理的。

然而，这并不是诸如“牛顿台球世界是不可计算的”断言的真正含义。我用来和我们的牛顿台球世界作比较的弗列得钦——托弗里“台球电脑”的特殊模型的确按照计算而进行。无论如何，这是弗列得钦和托弗里思想的基本点——他们模型的行为应该和一台（通用）电脑一样！我试图要提的问题是，在某种意义上，人类大脑驾驭适当的“不可计算的”物理定律，能比图灵机做得更好，这一点是不是可以理喻的。追究如下的问题将是徒劳的：

“如果球A永远碰不到球B，则你的问题的答案为‘非’。”

人们可能永远也等待不到断定问题中的球不会碰到一起的时刻！那当然正是图灵机行为的方式。

事实似乎很清楚地表明，牛顿台球世界在某个合适的意义上（至少在如果我们不管多碰撞的问题之时）是可计算的。人们通常计算这种行为的方法是做一些近似。我们可以想象这些球的中心被指定在点的网格上，譬如讲网格的点被划分到百分之几单位。时间也被认为是“离散”的，所有可允许的时刻是某一小单位（用Δt表示）的倍数。这就产生了使“速度”在一定程度上离散的可能性（两个连续允许的格点的位置的坐标值的差，除以Δt）。利用力的定律来计算加速度的适当的近似，再利用它使“速度”并因此下一允许时刻的新的格点位置被确定到所需要的近似程度。只要我们能维持所需的精度，则这种计算就可一直进行下去。很有可能算不了多少次其精度就失去了。以后的步骤是从更细的空间分格以及更细的时间间隔重新开始。这一回能得到更好的精度，并在

精度损失之前能计算到更久的将来的某一时刻。不断地增加细度，则计算的精度和所到达的将来的时间的长度就能不断地改进。可用这种方法将牛顿台球世界计算到任意高的精度（不管多碰撞的问题）——我们可以在这种意义上讲牛顿世界的确是可计算的。

然而，认为这一个世界在实际上是“不可计算的”断言是具有某种含义的。这是因为得知的初始数据的精度总是受限制的。这类问题的确存在着固有的不可忽视的“不稳定性”。初始数据的极为微小的改变会导致结果行为的绝大的变化。（任何玩台球的人，在他想用一个球去撞另一个球使之落入球囊时，都知道我这样说的意思！）这在（连续）碰撞发生时尤为明显。但是，这种不稳定性行为也会发生在牛顿的引力远距离作用时（多于两体的情况下）。所谓的“混沌”或“混沌行为”经常用来表示这种不稳定的类型。例如，混沌行为对天气影响重大。虽然我们对控制基本元素的牛顿方程式了解甚多，但是远期天气预报之不可靠性则是饱受诟病的！

这根本不是那类可以任何方式驾驭的“不可计算性”。这只是因为所知的初始态的精度有限，而终态不能由初态可靠地算出。实际上只是随机元素被引入到未来的行为中而已。如果大脑的确使用了物理定律中的不可计算性的有用的元素，则它们必须具有完全不同的，并从这里引出更正面得多的特性。相应地，我根本不把这种“混沌”行为称为“不可计算性”，而将其称为“不可预见性”。正如我们很快就会看到的，在（经典）物理学中的决定性的定律中存在不可预见性是一种非常一般的现象。在制造思维机器时，不可预见性正是我们希望尽量减小而不是去“驾驭”的东西！

为了更一般地讨论可计算性和不可预见性的问题，对物理定律采用比以前更广泛的观点将会更有帮助。这就促使我们不仅只考虑牛顿力学的理论，而且研究随后超越过它的各种理论。我们需要领略力学的美妙的哈密顿形式。

哈密顿力学

牛顿力学不仅在于非凡地应用到物理世界方面，而且在于它所引起的数学理论的丰富方面取得瞩目的成功。令人惊异的是，自然界所有超

等理论都被证明是数学观念的丰富来源。这些绝顶精确的理论就作为数学而言也是极富成果的，这个事实具有一种深刻和美丽的神秘。它毫无疑问地表明，在我们经验的实在世界和柏拉图的数学世界之间有某种根本的关联。（我将在第10章542页再讨论这些。）牛顿力学也许是这方面的一个顶峰，因为它一诞生即获得了微积分。而且，牛顿理论形成了非凡的称为经典力学的数学观念的实体。18世纪和19世纪许多伟大数学家的名字都和此发展相关联：欧拉、拉格朗日、拉普拉斯、刘维尔、泊松、雅科比、奥斯特罗格拉茨基、哈密顿。所谓“哈密顿理论”[8] 即为这一工作的总结。为了我们的目的对其稍微了解即可以了。威廉姆·罗曼·哈密顿（1805—1865）是一位多才多艺和富有创见的爱尔兰数学家，他还是在188页讨论过的哈密顿回路的发明者。他把力学发展成强调其与波传播相类似的形式。波和粒子的关系的暗示以及哈密顿方程的形式对于后来的量子力学的发展极为重要。我在下一章还会提及。

用以描述物理系统的“变量”是哈密顿理论的一个奇妙的部分。迄今为止，我们一直把粒子的位置当作基本的，而速度作为位置对时间的变化率。我们记得在牛顿系统中为了确定随后的行为，必须指定初始态（217页），也就是需要所有粒子的位置和速度。在哈密顿形式中，我们必须挑选粒子的动量，而不是速度。（我们在215页提到粒子动量是速度和质量的乘积。）这种改变似乎很微不足道，但是重要的在于每一粒子的位置和动量似乎被当作独立的量来处理。这样，人们首先“假装”不同粒子的动量和它所对应的位置的改变率没有什么关系，而仅仅是一组分开的变量。我们可以想象它们“可以”完全独立于位置的运动。现在在哈密顿形式中我们有两组方程。有一组告诉我们不同粒子的动量如何随时间变化，另一组告诉我们位置如何随时间变化。在每一种情况下，变化率总是由在该时刻的不同位置和动量所决定。

粗略地讲，第一组哈密顿方程表述了牛顿的关键的第二运动定律（动量变化率=力），而第二组方程告诉我们动量实际上即是依赖于速度（位置变化率=动量÷质量）。我们记得，伽利略——牛顿的运动定律是用加速度，即位置变化率之变化率（亦即“二阶”方程）来描述。现在，我们只需要讲到事物的变化率（“一阶”方程），而不是事物变化率的变化率。所有这些方程都是从一个重要的量推导而来：哈密顿函数H，它是系统的总能量按照所有位置和动量变量的表达式。

哈密顿形式提供了一种非常优雅而对称的力学描述，我们在下面写

出这些方程，仅仅是为了看看它们是什么样子的。虽然，甚至许多读者并不熟悉完全理解之所必需的微积分记号——它在这里是不需要的。就微积分而言，所有我们真正要理解的是，出现在每一个方程左边的点表示（在第一种情况下，动量的；在第二种情况下，位置的）对时间的变化率：

$$p_i = -\frac{\partial H}{\partial x_i}, \quad x_i = \frac{\partial H}{\partial p_i}$$

这里下标i用以区别所有不同的动量坐标p_1，p_2，p_3，p_4，……和所有不同的位置坐标x_1，x_2，x_3，x_4，……。n个不受限制的粒子具有3n个动量坐标和3n个位置坐标（每一个代表空间中的3个独立的方向）。符号∂表示“偏微分”（“在保持其他变量为常数的情况下取导数”）。正如前述的，H为哈密顿函数。（如果你不通晓“微分”，不必担心。只要认为这些方程的右边是某些定义完好的，以x_i 和p_i 来表达的数学式子就行了。）

实际上，坐标x_1 ，x_2 ，……和p_1 ，p_2 ，……可允许为某种比粒子通常的笛卡儿坐标（亦即x i为通常的沿三个不同的相互垂直的方向测量的距离）更一般的东西。例如坐标x i中的一些可以是角度（在这种情形下，相应的p_i 就是角动量，而不是动量，参见215页），或其他某些完全一般的测度。令人惊异的是，哈密顿方程的形状仍然完全一样。事实上，合适地选取H，哈密顿形式不仅仅是对于牛顿方程，而且对任何经典方程的系统仍然成立。对于我们很快就要讨论的麦克斯韦（—洛伦兹）理论，这一点尤其成立。哈密顿方程在狭义相对论中也成立。如果仔细一些，则广义相对论甚至也可并入到哈密顿框架中来。此外，我们将要看到在薛定谔方程（369页）中，哈密顿框架为量子力学提供了出发点。尽管1世纪以来发生的物理理论的所有革命性变化是如此地令人眼花缭乱，动力学方程结构的形式却是如此地统一，这真是令人惊叹！

相空间

哈密顿方程的形式允许我们以一种非常强而有力的一般方式去“摹想”经典系统的演化。想一个多维“空间”，每一维对应于一个坐标x_1，x_2，……，p_1，p_2，……（数学空间的维数，通常比3大得多。）此空间称之为相空间（图5.10）。对于n个无约束的粒子，相空间就有6n维（每个粒子有3个位置坐标和3个动量坐标）。读者或许会担心，甚至只要有单独一个粒子，其维数就是他或她通常所能摹想的2倍！不必为此沮丧！尽管6维的维数实在太多，很难画出，但是即使我们真的把它画出也无太多用处。仅仅就一满屋子的气体，其相空间的维数大约就有

10 000 000 000 000 000 000 000 000 000，

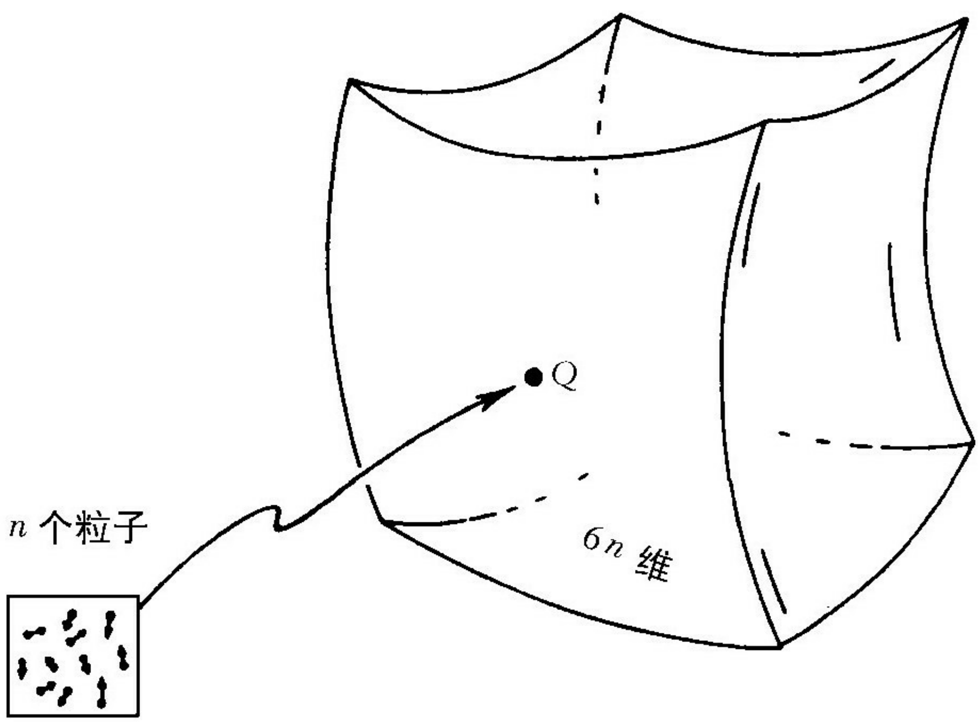


图5.10　相空间，相空间的单个点Q表明某一个物理系统的整个态，包括其所有部分的瞬态运动

去准确地摹想这么大的空间是没有什么希望的！既然这样，秘诀是甚至对于一个粒子的相空间都不企图去这样做。只要想想某种含糊的三维（或者甚至就只有二维）的区域，再看看图5.10就可以了。

我们如何按照相空间来摹想哈密顿方程呢？首先，我们要记住相空间

间的单个的点Q实际代表什么。它代表所有位置坐标x_1，x_2，……和所有动量坐标p_1，p_2，……的一种特定的值。也就是说，Q表示我们整个物理系统，指明组成它的所有单个粒子的特定的运动状态。当我们知道它们现在的值时，哈密顿方程告诉我们所有这些坐标的变化率是多少；亦即它控制所有单个粒子如何移动。翻译成相空间语言，该方程告诉我们，如果给定单个的点Q在相空间的现在位置的话，它将会如何移动。为了描述我们整个系统随时间的变化，我们在相空间的每一点都有一个小箭头——更准确地讲，一个矢量——它告诉我们Q移动的方式。这整体箭头的排列构成了所谓的矢量场（图5.11）。哈密顿方程就这样地在相空间中定义了一个矢量场。

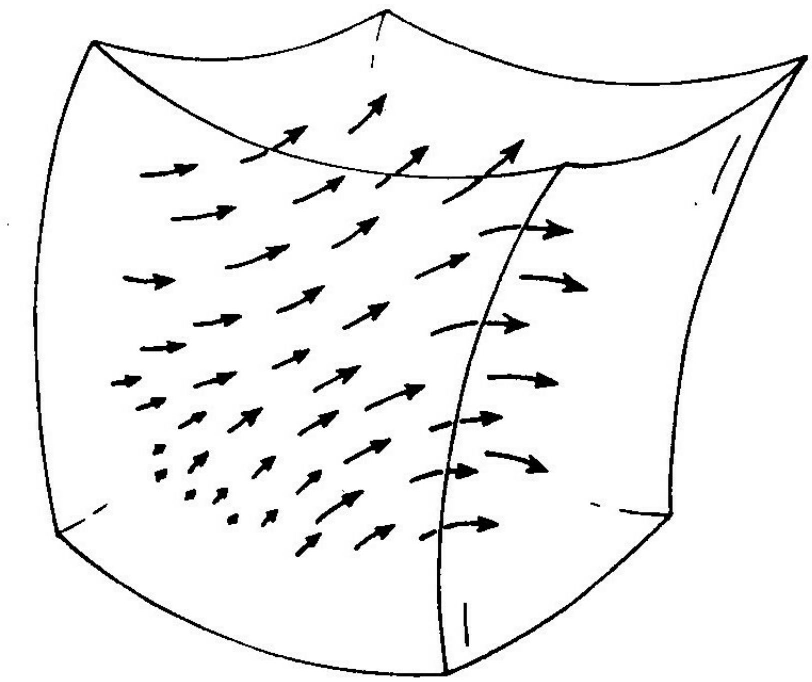


图5.11　相空间中的矢量场。它代表了按照哈密顿方程的时间演化

我们看看如何按照相空间来解释物理的决定论。对于时间t=0的初始数据，我们有了一组指明所有位置坐标和动量坐标的特定值；也就是说，我们在相空间特别选定了一点Q。为了找出此系统随时间的变化，我们就跟着箭头走好了。这样，不管一个系统如何复杂，该系统随时间的整个演化在相空间中仅仅被描述成一点沿着它所遭遇到的特定的箭头移动。我们可以认为箭头为点Q在相空间的“速度”。“长”的箭头表明Q移动得快，而“短”的箭头表明Q的运动停滞。只要看看Q以这种方式随

着箭头在时间t移动到何处，即能知道我们物理系统在该时刻的状态。很清楚，这是一个决定性的过程。Q移动的方式由哈密顿矢量场所完全决定。

关于可计算性又如何呢？如果我们从相空间中的一个可计算的点（亦即从一个其位置坐标和动量坐标都为可计算数的点，参阅第3章106页）出发，并且等待可计算的时间t，那么一定会终结于从t和初始数据计算得出的某一点吗？答案肯定是依赖于哈密顿函数H的选择，实际上，在H中会出现一些物理常量，诸如牛顿的引力常量或光速——这些量的准确值视单位的选定而被决定，但其他的量可以是纯粹数字——并且，如果人们希望得到肯定答案的话，则必须保证这些常量是可计算的数。如果假定是这种情形，那我的猜想是，答案会是肯定的。这仅仅是一个猜测。然而，这是一个有趣的问题，我希望以后能进一步考察之。

另一方面，由于类似于我在讨论有关台球世界时简要提出的理由，对我来说，这似乎不完全是相关的问题。为了使一个相空间的点是不可计算的断言有意义，它要求无限精确的坐标——亦即它的所有小数位！（一个由有尽小数描述的数总是可以计算的。）一个数的小数展开的有限段不能告诉我们任何关于这个数整个展开的可计算性。但是，所有物理测量的精度都是有限的，只能给出有限位小数点的信息。在进行物理测量时，这是否使“可计算数”的整个概念化成泡影？

的确，一个以任何有用的方式利用某些物理定律中（假想的）不可计算因素的仪器不应依赖于无限精确的测量。也许我在这里有些过分苛刻了。假定我们有一台物理仪器，为了已知的理论原因，模拟某种有趣的非算法的数学过程。如果此仪器的行为总可以被精密地确定的话，则它的行为就会给一系列数学上有趣的没有算法的（像在第4章中考虑过的那些）是/非问题以正确答案。任何给定的算法都会到某个阶段失效。而在那个阶段，该仪器会告诉我们某些新的东西。该仪器也许的确能把某些物理常量测量到越来越高的精度。而为了研究一系列越来越深入的问题，这是需要的。然而，在该仪器的有限的精度阶段，至少直到我们对这系列问题找到一个改善的算法之前，我们得到某些新的东西。然而，为了得到某些使用改善了的算法也不能告诉我们的东西，就必须乞求更高的精度。

尽管如此，不断提高物理常量的精度看来仍是一个棘手和不尽如人

意的信息编码的方法。以一种离散（或“数字”）形式得到信息则好得多。如果考察越来越多的离散单元，也可重复考察离散单元的固定集合，使得所需的无限的信息散开在越来越长的时间间隔里，因此能够回答越来越深入的问题。（我们可以将这些离散单元想象成由许多部分组成，每一部分有“开”和“关”两种状态，正如在第2章描述的图灵机的0和1状态一样。）由此看来我们需要某种仪器，它能够（可区别地）接纳离散态，并在系统按照动力学定律演化后，又能再次接纳一个离散态集合中的一个态。如果事情是这样的话，则我们可以不必在任意高的精度上考察每一台仪器。

那么，哈密顿系统的行为确实如此吗？某种行为的稳定性是必须的，这样才能清晰地确定我们的仪器实际上处于何种离散态。一旦它处于某状态，我们就要它停在那里（至少一段相当长的时间），并且不能从此状态滑到另一状态。不但如此，如果该系统不是很准确地到达这些状态，我们不要让这种不准确性累积起来；我们十分需要这种不准确性随时间越变越小。我们现在设想的仪器必须由粒子（或其他子元件）所构成。需要以连续参数来描述粒子，而每一个可区别的“离散”态图覆盖连续参数的某个范围。（例如，让粒子停留在2个盒子中的一个便是一种表达离散双态的方法。为了指明该粒子确实是在某一个盒子中，我们必须断定其位置坐标在某个范围之内。）用相空间的语言讲，这表明我们的每一个“离散”的态必须对应于相空间的一个“区域”，同一区域的相空间点就对应于我们仪器的这些可选择的同一态（图5.12）。

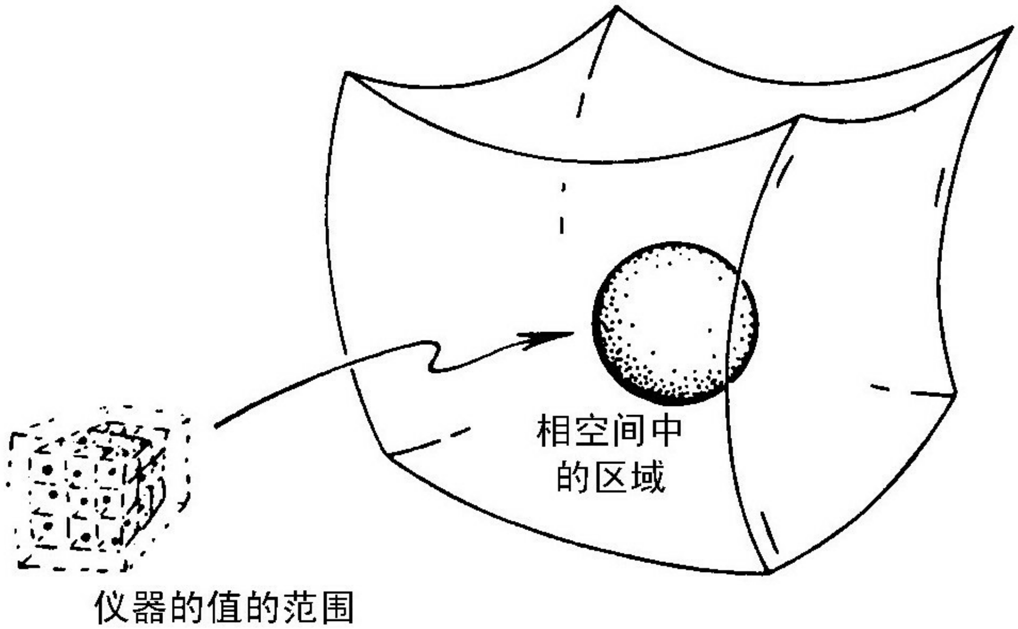


图5.12　相空间中的一个区域对应于所有粒子位置和动量的可能值的一个范围。这样的区域可代表某仪器一个可区别态（亦即“选择”）

现在假定仪器在开始时的态对应于它的相空间中的某一个范围R_0。我们想象R_0 随着时间沿着哈密顿矢量场被拖动，到时刻t该区域变成R_t。在画图时，我们同时想象对应于同一选择的所有可能的态的时间演化（图5.13）。关于稳定性的问题（在我们感兴趣的意义上讲）是，当t增加时区域R t是否仍然是定域性的，或者它是否会向相空间散开去。如果这样的区域在时间推进时仍是定域性的，我们对此系统就有了稳定性的量度。在相空间中相互靠近的点（这样它们对应于相互类似的系统的细致的物理态）将继续靠得很近，给定的态的不准确性不随时间而放大。任何不正常的弥散都会导致系统行为的等效的非预测性。

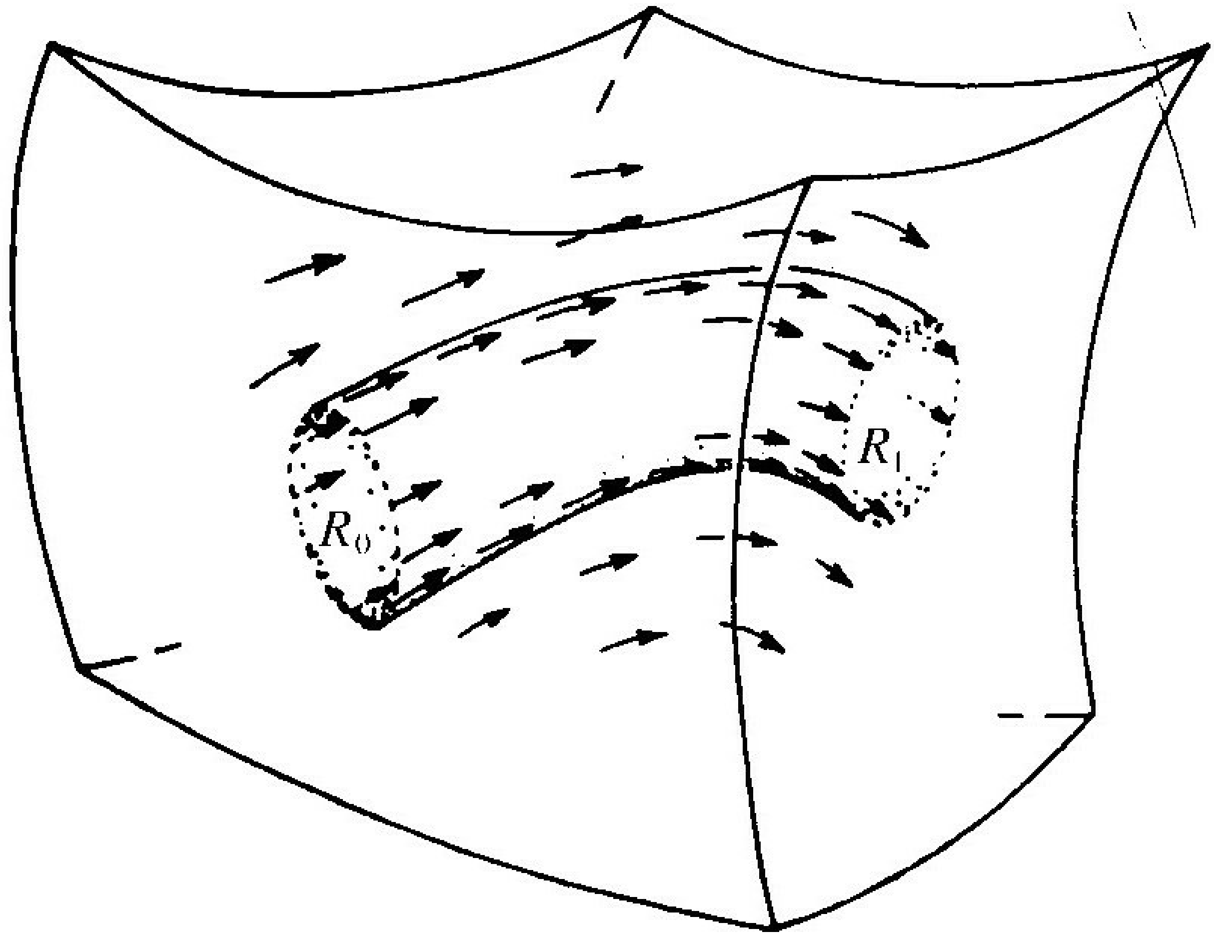


图5.13　随着时间演化，相态区域R_0 沿着矢量场被拖到一个新区域R_t 。这可表示我们仪器的某一特定选择的时间演化

我们对于哈密顿系统一般地可以说什么呢？相空间的区域究竟是否随时间散开呢？似乎对于一个如此广泛的问题，很少有什么可说的。然而，人们发现了一个非常漂亮的定理，它要归功于杰出的法国数学家约瑟夫·刘维尔（1809—1882）。该定理讲，相空间中的任何区域的体积在任何哈密顿演化下必须保持常数。（当然，由于我们的相空间是高维的，所以“体积”必须是在相应高维意义上来说的。）这样，每一个R_t的体积必须和原先的R_0的体积一样。初看起来，这给了我们的稳定性问题以肯定的答案。在相空间体积的这层意义上，我们区域的尺度不能变大，好像我们的区域在相空间中不会散开似的。

然而，这是使人误解的。我们在深思熟虑之后就会感到，很可能情况刚好与此相反！在图5.14中我想表示人们一般预料到的那种行为。我

们可以将初始区域R_0想象成一个小的、“合理的”，亦即较圆的而不是细长的形状。这表明属于R_0 的态在某种方面不必赋予不合情理的精确性。然而，随着时间的发展，区域R_1 开始变形并拉长——初看起来有点像变形虫，然后伸长到相空间中很远的地方，并以非常复杂的方式纠缠得乱七八糟。体积的确是保持不变，但这个同样小的体积会变得非常细，再发散到相空间的巨大区域中去。这和将一小滴墨水放到一大盆水中的情形有点类似。虽然墨水物质的实际体积不变，它最终被稀释到整个容器的容积中去。区域R_t在相空间中的行为与此很类似。它可能不在全部相空间中散开（那是称之为“遍历”的极端情况），但很可能散开到比原先大得极多的区域去。（可参阅*Davies 1974* 的进一步讨论。）

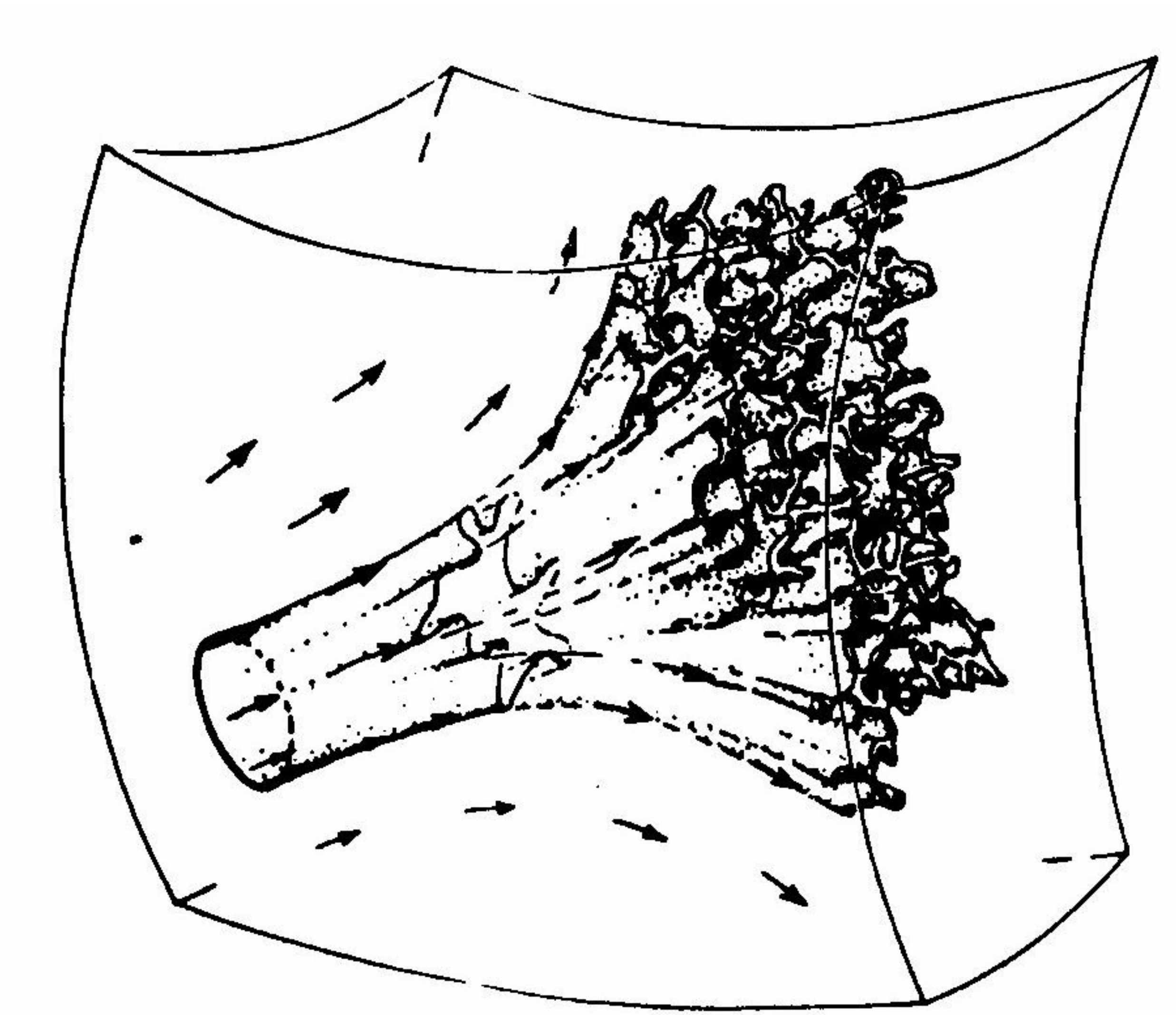


图5.14　尽管刘维尔定理告诉我们，随着时间演化相空间体积不变，但是由于该演化的极端复杂性，这个体积通常会等效地弥散开来

麻烦在于保持体积并不意味就保持形状：小区域会被变形，这种变形在大距离下被放大。由于在高维时存在区域可以散开去的多得多的“方向”，所以这问题比在低维下严重得多。事实上，刘维尔定理远非“帮助”我们将区域R t控制住，而是向我们提出了一个基本问题！若无刘维尔定理，我们可以摹想相空间中区域的毫无疑义的发散趋势可由整个空间的缩小而补偿。然而，这一个定理告诉我们这是不可能的，而我们必须面对这个惊人的含义——这个所有正常类型的经典动力学（哈密顿）系统的普适的特征[9]！

鉴于这种发散到整个相空间去的行为，我们会问，经典力学怎么可能作出预言？这的确是一个好问题。这种弥散所告诉我们的是，不管我们多么精确地（在某一合理的极限内）知道系统的初始态，其不确定性将随着时间而不断增大，而我们原始的信息几乎会变得毫无用处。在这个意义上讲，经典力学基本上是不可预言的。（回想前面考虑过的“混沌”概念。）

那么，何以迄今为止牛顿动力学显得如此之成功呢？在天体力学中（亦即在引力作用下的天体）其原因在于，第一，有关的凝聚的物体数目相对很少（太阳、行星和月亮），这些物体的质量相差悬殊——这样在估量近似值时，可以不必管质量更小物体的微扰效应，而处理更大的物体时，仅仅需要考虑它们相互作用的影响；第二，可以看到，适用于构成这些物体的个别粒子的动力学定律，也可以在这些物体本身上的水平上适用——这使得在非常好的近似下，太阳、行星和月亮实际上可以当作粒子来处理，我们不必去为构成天体的单独粒子的运动的微小细节担忧[10]。我们再次只要考虑“很少”的物体，其在相空间中的弥散不重要。

除了天体力学和投掷物行为（它其实是天体力学的一个特例）之外，只牵涉到小数目的粒子的简单系统的研究，牛顿力学所用的主要方法是根本不管这些细节的“可决定性的预言的”方面。相反地，人们利用一般的牛顿理论做模型，从这些模型可以推导出整体行为。某些诸如能量、动量和角动量守恒定律的准确推论的确在任何尺度下都有效。此外，存在可与制约单独粒子的动力学规律相结合的统计性质，它能对有关的行为作总体预言。（参阅第7章关于热力学的讨论；我们刚讨论过的相空间弥散效应和热力学第二定律有紧密的关系。我们只要相当仔细，便可利用这些观念作预言。）牛顿本人所做的空气声速的计算（1

个世纪后拉普拉斯进行了微小的修正）便是一个好例子。然而，牛顿（或更笼统来说，哈密顿）动力学中固有的决定性在实际中适用的机会非常稀少。

相空间弥散效应还有一个惊人的含义。它告诉我们，经典力学不能真正地描述我们的世界！我说得有点过分了一些，但是并不太过分。经典力学可以很好地适用于流体——特别是气体的行为，在很大的程度上适用于液体——此处人们只关心粒子系统的“平均”性质，但是在对固体作计算时就出了毛病，这里要求知道更细节的组织结构。固体由亿万颗点状的粒子所组成，由于相空间弥散其排列的有序性应不断地降低，何以保持其形状大致不变呢？正如我们已经知道的，量子力学在理解固体的实在结构时是不可或缺的。量子效应可多多少少防止相空间的弥散（第8章和第9章）。

这也和制造“计算机器”的问题相关。相空间弥散是某种必须控制的东西。相空间中对应于一个电脑的“离散”态的区域（例如前述的R_0）不应允许其过度弥散开来。我们记得，甚至弗列得钦——托弗里“台球电脑”需要某种外围的固体墙才能工作。包括许多粒子的物体的“刚性”正是需要量子力学起作用的某种东西。看来，甚至“经典”电脑也必须借助于量子物理学的效应才能有效地工作！

麦克斯韦电磁理论

在牛顿的世界图像中，人们设想一个微小粒子靠一种超距作用的力作用到另一个粒子上。如果粒子不是完全点状的，可以认为由于偶尔的实际物理接触而互相反弹离开。正如我前面（218页）提到的，电学和磁学（古人即知道此两者的存在，威廉·吉尔伯特在1600年和本杰明·富兰克林在1752年分别进行了一些细节的研究）的行为和引力很类似。虽然同号的电荷（磁极强度）相互排斥而不是吸引，它们都以距离的平方反比律衰减。这里的电磁力是由电荷（磁极强度），而不是由质量决定其强度。在这个水平上，将电学和磁学归并到牛顿理论中去并没有什么困难。光的行为也可以粗略地（虽然有某些困难）容纳进去。我们或者将光当作单独粒子（正如我们现在应称之为“光子”的那样）组成，或者把它当作某种媒质中的波的运动。在后一情况该媒质（“以太”）本身应认为是由粒子组成的。

运动电荷会产生磁力的这一事实引起了额外的复杂性，但是这并没有把整个体系瓦解。大量的数学家和物理学家（包括高斯）提出了在一般牛顿框架中似乎满意的、描述运动电荷效应的方程组。第一位向这个“牛顿式”的图像提出严肃挑战的科学家是英国伟大的实验家兼理论家米凯尔·法拉第（1791—1867）。

为了理解这个挑战的性质，我们首先要定义物理场的概念。首先考虑磁场。大部分读者都有过这样的经验，将一张纸放在磁铁上时，纸上的铁粉末具有特别的形态。这些粉末以一种令人惊异的方式沿着所谓的“磁力线”串起来。我们可以想象，即便粉末不在该处，磁力线仍在那里。它们构成了我们称之为磁场的东西。这“场”在空间的每一点都朝着一定的方向，亦即在该点力线的方向。实际上，我们在每一点都有一个矢量。这样，磁场就给我们提供了一个矢量场的例子。（我们可把它和上一节考虑的哈密顿矢量场相比较，但现在这一个矢量场是在通常的空间中，而不在相空间中。）类似地，一个带电的物体被一种称之为电场的不同种类的场所围绕；而且引力场也类似地围绕着任何有质量的物体。这些也都是空间的矢量场。

远在法拉第之前，人们就有了这些观念，它们已成为牛顿力学理论家的一部分武器。但是认为这种“场”中不包含实际物理物质的观点占优势。反之，它们被当作为某一个粒子放在不同的点时所作用的力提供一种必要的“簿记”。然而，法拉第深刻的实验发现（利用运动线圈、磁铁等）使他坚信，电磁场是真正的“东西”，并且变化的电磁场有时会相互“排挤”到原先空虚的空间，以产生一种脱离物体的波动！他猜测到光也许就包括这类波动。这种观点背离了占统治地位的“牛顿智慧”。按照牛顿的观点，这类场不能在任何意义上被认为是“真实的”，而仅仅是作为“真正的”牛顿点粒子超距作用“实在”图像的方便的数学辅助物而已。

面临着法拉第以及优秀的法国物理学家安德列·玛雷·安培（1775—1836）和其他人更早的实验发现，伟大的苏格兰物理学家兼数学家詹姆斯·克拉克·麦克斯韦（1831—1879）对从这些发现产生的电磁场方程的数学形式感到疑惑。他以惊人的灵感，对这些方程作了初看起来似乎非常微小的，但却是含义深远的改变。这个改变根本不是由已知的实验事实（虽然与之相协调）暗示的。这是麦克斯韦理论自身所要求的结果，部分是物理学上的，部分是数学上的，还有部分是美学上的。麦克斯韦方程的一个含义是电磁场的确在空虚的空间中相互“推挤”。振荡的磁场

产生振荡的电场（这是法拉第的实验发现所隐含的）。而振荡的电场又反过来产生振荡的磁场（由麦克斯韦理论推导得来的），并且这又接着产生电场，等等。（这种波的详图见344页的图6.26和346页的图6.27。）麦克斯韦能够算出这种效应在空间传播的速率——并且他发现这正是光的速率！此外，这些所谓的电磁波还展示出了很久以来就知道的干涉和令人困惑的偏振性质（我们在第6章298页、344页还要回到这些上来）。除了说明波长在一个特定范围（4×10^{-7} ～7×10^{-7} 米）的可见光的性质外，还预言了导线中电流产生的其他波长的电磁波。出色的德国物理学家亨利希·赫兹于1888年在实验上证实了这种波的存在。法拉第的富有灵感的希望在美妙的麦克斯韦方程中的确找到了坚实的基础！

虽然我们在这儿并不必了解麦克斯韦方程的细节，稍微看看它们是什么样子并没有什么害处：

$$\frac{1}{c^2}\cdot\frac{\partial \boldsymbol{E}}{\partial t}=\operatorname{curl}\boldsymbol{B}-4\pi\boldsymbol{j},\quad \frac{\partial \boldsymbol{B}}{\partial t}=-\operatorname{curl}\boldsymbol{E}$$

$$\operatorname{div}\boldsymbol{E}=4\pi\rho,\qquad \operatorname{div}\boldsymbol{B}=0.$$

此处E、B和j分别为电场、磁场和电流；ρ为电荷密度，c只是一个常数，也就是光速[11]。不必忧虑curl及div等项，它们简单地表示不同类型的空间变化。（它们是某种相对于空间坐标的偏微分算符的组合。可以回想我们在讨论哈密顿方程时遇到的用符号∂表示的偏微分运算。）在前面两个方程左边出现的算符∂/∂t实际上和用在哈密顿方程的点一样，其不同之处只是技术性的。这样∂E/∂t表示电场的变化率，而∂B/∂t表示磁场的变化率。第一个方程[38]说明电场如何按照磁场和电流在该时刻的行为而变化；而第二个方程说明磁场如何按照电场在该时刻的行为而变化。第三个方程粗略地讲是平方反比律的另一种形式，它是讲（该时刻的）电场必须和电荷分布相关；而第四个方程是对磁场说同样的东西，除了在这情况下没有“磁荷”（或分开的“北极”或“南极”粒子）以外。

这些方程在下面这一点和哈密顿的很相像，即依据在任何给定时刻的电场和磁场的值，它们给出了这些量对时间的变化率。所以麦克斯韦方程和通常的哈密顿理论一样是决定论的。仅有的也是一个重要的差别是，麦克斯韦方程是场方程而不是粒子方程。这表明我们需要用无穷多个参数去描述系统的态（空间中的每一点的场矢量），而不仅仅需要像在粒子论中的有限的数目参数（每个粒子的3个位置和3个动量坐标）。因此麦克斯韦理论的相空间是无限维的！（正如我以前提到过的，一般的哈密顿框架，实际上可以包容麦克斯韦方程。但由于这无限的维数，该框架必须稍微推广一下[12]。）

麦克斯韦理论为我们的物理实在的图像添加上具有根本性的新的部分。我们必须接受场自身的存在，而不能把它仅仅当作牛顿物理中的“实在”粒子的数学的附属物。在这一点上它超越了我们的原先的理论框架。麦克斯韦的确向我们指出，当场以电磁波传播时，它们自身携带一定量的能量。他还给出了这种能量的显明的表达式。从一处传播到另一处的“脱离物体”的电磁波能传递能量的这一惊人事实，最终由赫兹在实验上探测到它的存在而被证实。这个事实虽然如此惊人，而现在却变成这么熟悉的东西了。

可计算性和波动方程

麦克斯韦能直接从他的方程推导出，在没有电荷或电流（亦即在上述方程中j=0，ρ=0）的空间区域，所有电磁场的分量必须满足一个称为波动方程[39]的方程。由于波动方程是关于一个单独的量的，而不是电磁场的所有6个分量的方程，所以可视作麦克斯韦方程的“简写”。它的解表现了类似波动的行为，并牵涉到诸如麦克斯韦理论的“极化”（电场矢量的方向，见344页）等其他复杂性。

因为波动方程及其可计算性的关系已被清楚地研究过，所以我们对它格外有兴趣。事实上，玛利安·玻依堪·玻——埃勒和因·里查德（1979，1981，1982和1985）指出，尽管波动方程在平常的意义上具有决定性的行为——亦即初态数据一被提供，则其他时刻的解即被决定——还存在某种古怪类型的可计算的初始数据，它使得在以后可计算的时刻被决定的场的值实际上是不可计算的。这样，此一似是而非的物理场论的方程（虽然不完全是在我们世界中实际成立的麦克斯韦方程）会

在玻——埃勒和里查德的意义上产生不可计算的演化！

这结果在表面上似乎相当令人震惊——这看来和我在上一节的猜测相抵触，除了那时人们关心的是“合理的”哈密顿系统的可能的可计算性以外。然而，玻——埃勒和里查德结果固然是惊人的并和数学有关系，它和猜测的冲突并没有什么真正的物理意义。原因在于，他们“古怪”的初始数据不以一种通常人们对物理上有意义的场所要求的方式而“光滑地改变”[13]。玻——埃勒和里查德实际上证明了，如果我们不容许这一类场，则不会产生不可计算性。无论如何，甚至如果允许这类场，很难想象任何物理“仪器”（诸如人脑？）能利用这样的“不可计算性”。这只有当允许作任意高精度的测量时才相干。但正如我说过的，这在物理上不是非常现实的。尽管如此，玻——埃勒和里查德的结果代表了一个重要研究领域的美妙开端，迄今这个领域还很少被研究过。

洛伦兹运动方程；逃逸粒子

麦克斯韦方程本身还不是一个完整的方程组。如果给定了电荷和电流的分布，则它们提供了电磁场传播方式的美妙的描述。在物理上，这些电荷主要是我们知道的电子和质子等带电粒子，而电流是由这种粒子的运动所引起的。如果我们知道这些粒子在何处并如何运动，则麦克斯韦方程告诉我们电磁场会如何行为。该方程并没有告诉我们这些粒子自身如何行为，此问题的部分答案在麦克斯韦年代即已经知道，但直到1895年杰出的荷兰物理学家亨德里克·安东·洛伦兹利用与狭义相对论有关的思想去推导现在称之为带电粒子的洛伦兹运动方程后（参阅Whittaker　1910，P310，P395），才得到令人满意的方程组。这些方程告诉我们带电粒子的速度如何因所处的电磁场的影响而连续地改变[14]。把洛伦兹方程和麦克斯韦方程相联立，人们便能同时得到带电粒子的电磁场的时间演化的规则。

然而，这一套方程并非一切都相安无事。如果一直到粒子的自身的直径的尺度之下（电子的“经典半径”大约为10-15米）场都是非常均匀的，而且粒子运动也不过分激烈的话，则它们给出了极好的结果。但此处存在一个原则上的困难，在其他情况下它会变得重要起来。洛伦兹方程要我们去做的是考察带电粒子所在处的准确的那一点的电磁场（并且实际上提供了该点的“力”）。如果粒子是有限尺度的，则那一点应如何

选取呢？是否我们应取粒子的“中心”，或是对表面上所有点的场（“力”）取平均？如果场在粒子尺度下不是均匀的，则这就产生了差异。还有更严重的问题：粒子表面（或中心）的场究竟如何？记住我们考虑的是一个带电的粒子。粒子本身引起的电磁场必须叠加到粒子所处的地方的“背景场”上去。粒子的自身场在靠近“表面”处变得极强，并且轻而易举地糟蹋它附近的所有其他的场。而且，围绕着自身的粒子场会多多少少地指向外面（或内面）。这样粒子所要反应的总的实际场根本不是均匀的，在粒子“表面”的不同地方指向不同的方向，更不要说它的“内部”了（图5.15）。现在我们必须开始忧虑，互异的作用到粒子上的力是否使之旋转或变形，我们必须知道它的弹性性质等（并且这里还有一个和相对论有关的特别有疑问的问题，我先不在此烦恼读者）。显然，这个问题比初看时复杂得多。

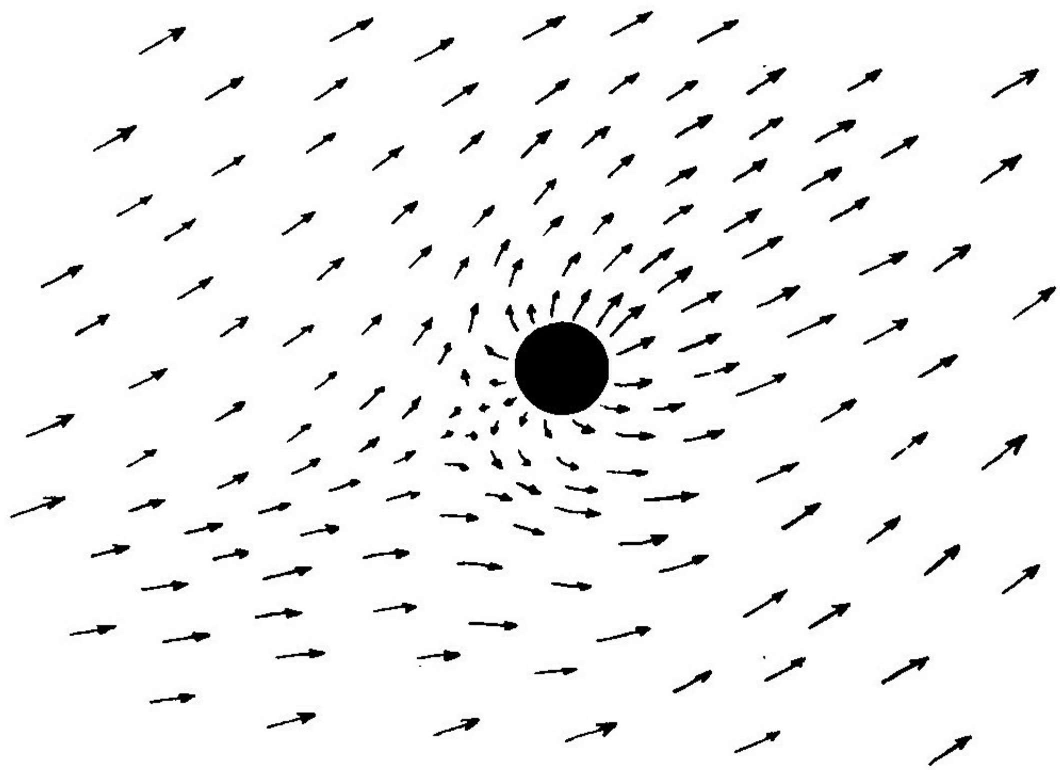


图5.15　我们要如何严格地应用洛伦兹运动方程？由于自己的场在粒子的位置处起主导作用，作用在它上面的力不能简单地从该粒子所处地方的场得到

也许我从一开始就把粒子当做点粒子会更好些。但这会导致另一类问题，因为在粒子的邻近处其自身的电场会变成无穷大。按照洛伦兹方程，如果它必须对它所处的地方的电磁场响应，则它必须对此无穷大的

场响应！为了使洛伦兹力定律有意义，必须找出一种减去粒子自身的场以剩下有限的背景场的方法，这样粒子才能毫不含糊地对背景场响应。1938年狄拉克（我们在后面还要提到他）解决了这个问题。但是，狄拉克解导出了某些令人恐慌的结论。他发现为了决定粒子和场的行为，不但必须知道每个粒子的初始位置和速度，也必须知道其初始加速度（这是一种在标准的动力学理论的范围内不太正常的情况）。对大多数的初始加速度值，粒子的最终行为变得完全疯狂，它自发地加速并很快地趋近于光速！这就是狄拉克的“脱逸解”，它并不对应于任何实际发生在自然里的东西。人们必须找到一种正确选择初始加速度以避免脱逸解的方法。只有一个人使用“先知”——也就是，必须指明能最终导出脱逸解的初始加速度并避免之，才能做到。这根本就不是在一个标准的决定性的物理问题中选择初始条件的方法。在传统的决定论中，这些初始数据可以任意给定，不受任何未来的行为要求的约束。而在这里，不仅是将来完全决定了在过去某一时刻的应选取的初始值，而且这些非常特别的数据由于要使未来行为确实“合理”的要求，而被非常苛刻地约束。

基本的经典方程就只能走到这么远。读者会意识到经典物理定律中决定性和可计算性的问题真是乱麻一团，在物理学定律中是否有一个目的论的因素呢？未来是否对过去允许发生的事有某种影响呢？实际上，物理学家并未认真地将这些经典电动力学（经典带电粒子和电磁场的理论）的含义当作实在的描述。他们对上述困难的通常回答是，带电的单独粒子问题是在量子电动力学范畴里，我们不能指望利用纯粹经典过程得到有意义的答案。这无疑是对的。但正如我们以后将要看到的，在这一点上量子理论自身也有问题。事实是，狄拉克正是因为想到，也许能为解决（物理上更适当的）量子问题中的甚至更大的基本困难得到灵感，而考虑带电粒子的经典问题。以后我们必须面临量子理论的这个问题！

爱因斯坦和庞加莱狭义相对论

我们回顾一下伽利略的相对性原理。它告诉我们，如果我们从一个静止坐标系转换到运动坐标系，伽利略和牛顿的物理定律完全不变。这意味着仅仅考察在我们周围的物体的动力学行为，不能确定我们是处于静止状态，还是沿着某一方向作匀速运动。（回忆一下211页至212页描述伽利略在海上的船。）当我们将麦克斯韦方程合并到这些定律中去

时，伽利略的相对论仍然对吗？我们知道麦克斯韦电磁波以固定的速率——即光速传播。常识似乎告诉我们，如果我们在某一方向非常快地运动，则光在那一方向相对我们的速率应减少到比c小（因为我们沿着那个方向去“追逐”光线），而且在相反的方向光速应相应地增加到比c大（因为我们向着光运动）——这都和麦克斯韦理论的不变的值c不一致。确实，常识似乎是对的：合并的牛顿和麦克斯韦方程不满足伽利略相对论。

正是由于对这个问题的忧虑导致爱因斯坦于1905年——事实上庞加莱在他之前（1898—1905）——提出狭义相对论。庞加莱和爱因斯坦各自独立地发现麦克斯韦方程也满足一个相对性原理（参阅*Pais 1982*）；也就是如果我们从一个静止坐标系换到运动坐标系时，方程也有类似的不变的性质。虽然在这种情况下，变换规则和伽利略——牛顿物理不相容！为了使两者相容，必须修正其中的一组方程——或者抛弃相对性原理。

爱因斯坦不想抛弃相对性原理。他凭着超等的物理直觉坚持，这个原则必须对于我们世界的物理定律成立。此外，他知道伽利略——牛顿物理对于所有的已知现象，只在速度和光速相比很微小的情况下被检验，这时不相容性并不显著。而人们早已知道，只有光本身的速度很大，才足以使这种偏离变得重要。所以，正是光的行为才能告诉我们究竟要采用何种相对性原理——而制约光的方程正是麦克斯韦方程。这样适合于麦克斯韦理论的相对性原理要保留；而相应地伽利略——牛顿定律要作修正！

在庞加莱和爱因斯坦之前，洛伦兹也致力于解决并部分回答了问题，直到1895年，洛伦兹采取的观点认为将物质结合在一起的力具有电磁性（后来证明正是如此）。这样，实在物体的行为应该满足从麦克斯韦方程推导出的定律。其中一个推论，是以与光速可相比拟的速度运动的物体在运动的方向会有微小的收缩（所谓的“费兹杰拉德——洛伦兹收缩”）。洛伦兹利用它来解释迈克耳孙和莫雷在1887年进行的令人困惑的实验发现。该实验似乎指出不能用电磁现象来确定一个“绝对”静止的坐标系。（迈克耳孙和莫雷指出，地球表面上的光的表观速度不受地球绕太阳公转的影响，这和预想的非常不一样。）是否物体的行为总是这样，以至于不可能在局部检验它的匀速运动呢？这是洛伦兹的近似的结论；而且他只局限于物体的特殊的，也就是认为只有电磁力才有意义

的理论。作为一位杰出的数学家，庞加莱在1905年指出，麦克斯韦方程基础的相对性原理，物体有一个精确的行为方式使得局部检测物体的匀速运动根本办不到。他并透彻地了解了此原理的物理含义（包括我们很快就要考虑到的“同时性的相对性”）。他认为这仅仅是一种可能性，而不像爱因斯坦那样坚持相对性原理必须成立。

麦克斯韦方程满足的相对性原理后来被称作狭义相对论。要掌握它不甚容易。它有许多反直观的特征，一下子很难把这些特征当作我们生活其中的世界的性质接受下来。事实上，若不是富有创见和洞察力的俄国/德国几何学家赫曼·闵可夫斯基（1864—1909）于1908年引进了进一步的要素，很难对狭义相对论赋予意义。闵可夫斯基曾是爱因斯坦在苏黎世高等理工学院的导师。1908年，闵可夫斯基在他发表在格丁根大学的著名演讲中说道：

从今以后空间自身以及时间自身必像影子般地渐渐消退，只有两者的某种结合保持为独立的实体。

现在，让我们按照美妙的闵可夫斯基空间——时间来理解狭义相对论的基础。

和空间——时间概念相关的一个困难在于它是四维的，这样要去摹想它就非常困难。然而，我们已逃过了相空间这一关，区区四维不会引起我们太多的麻烦！和以前一样，我们将采用“欺骗”的手法把空间画成更少的维数——但是，这回欺骗的程度没有过去那么严重，我们的图画也相应地更为准确一些。二维图（一维空间和一维时间）对许多目的是足够的。但我还希望读者允许我有点更冒险地升高到三维图（二维空间和一维时间）。这样子我们就得到了非常好的图画，并在原则上认为不必做许多改变就可将三维图的观念推广到四维的情况去。关于空间——时间图要记住的是，在它上面的每一点代表一个事件——也就是某一时刻的空间的一点，只有瞬息存在的一点。整个图代表过去、现在和将来的全部历史。因为一个粒子总存留在时间内，所以它不是以一点，而是以称作粒子的世界线的一条线来代表。如果粒子做匀速直线运动，则其世界线为直线。如果它做加速运动（亦即非匀速运动），则世界线是弯曲的。世界线描述了粒子存在的整个历史。

我在图5.16中画出了具有二维空间和一维时间的空间——时间图。

我们可想象沿着垂直方向测量有一标准的时间坐标t，以及在水平方向测量的两个空间坐标x/c和z/c [40] ，在中心处的圆锥是空间——时间原点O的（未来）光锥。为了领略其意义，可以想象在事件O处发生一次爆炸。（此爆炸在时刻t=0发生在空间的原点。）从爆炸发出的光的历史正是此光锥。在二维空间中看，闪光是以基本的光速c向外运动的圆圈。在全部三维空间中看，变成以光速c向外运动的一个球面——光的波前的球面——但是我们在这儿压缩了空间方向，所以只得到了一个圆圈，正如从一块石头落到水池中去的那一点发出的涟漪的圆圈那样。如果我们在向上的方向连续截割光锥的话，就能在此空间——时间中看到这一圆圈。这些水平面代表随时间坐标t增加时不同的空间的描述。相对论的一个特征是，一个物质粒子不能以比光速更快的速度运动（后面还要讲到）。所以从爆炸出来的物质粒子必须落到闪光的后头。用空间——时间的语言来说，这表明所有这些粒子的世界线必须在光锥内部。

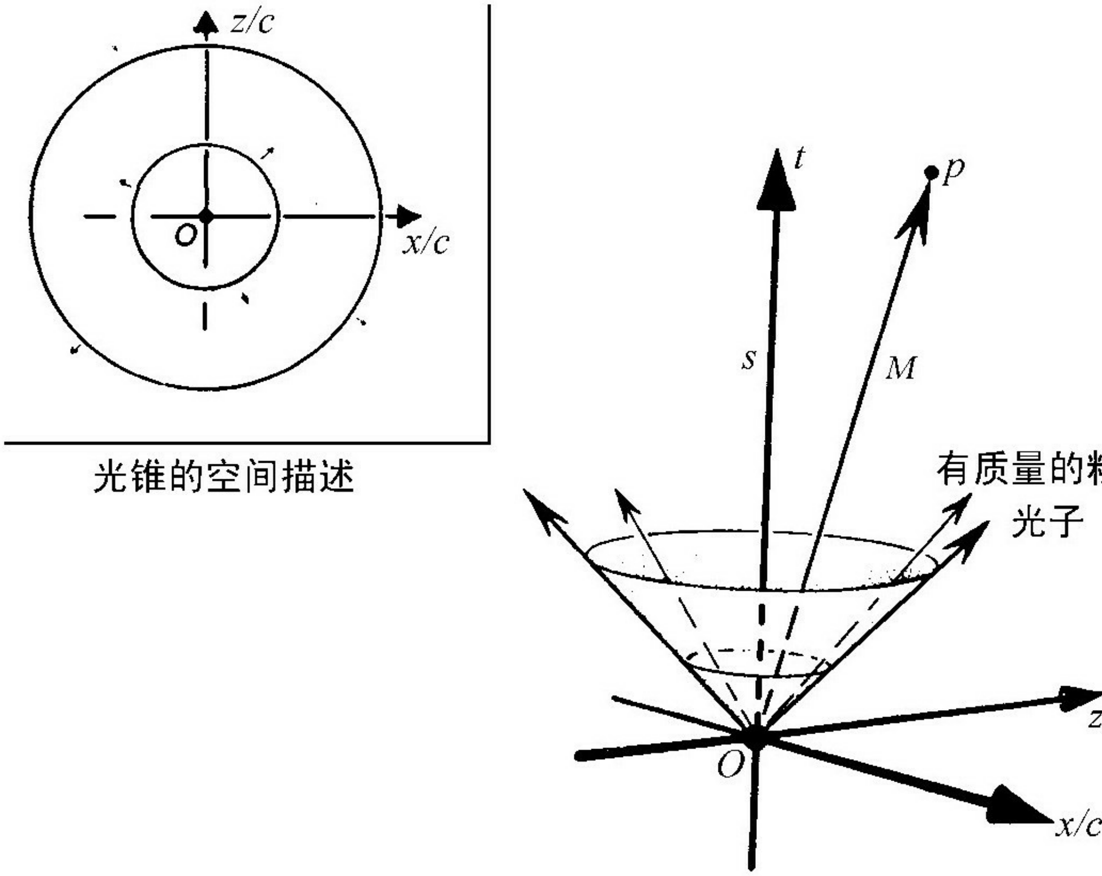


图5.16　闵可夫斯基空间——时间（仅有两维空间）中的一个光锥，描述了在空间——时间原点的事件O处发生的爆炸的闪光的历史

用称作光子的粒子比用电磁波来描述光更为方便。此刻我们暂时可以将一个“光子”当作一个电磁场高频振动的小“波包”。在下一章我们将要讨论的量子描述中，这个术语的物理意义将会更清楚。但在这里“经典”光子对我们也是有用的。在自由空间中光子总是以基本速度c沿直线运动。这表明在闵可夫斯基空间——时间图中光子的世界线总是画成一条和垂直线倾斜45°的直线。在O点处的爆炸产生的光子描写了一个中心位于O的光锥。这些性质在空间——时间的所有点都应成立。原点并没有任何特别之处；点O和任何其他点无区别。这样的空间——时间的每一点都必须有一个和在原点光锥具有同样意义的光锥。如果我们宁愿使用光的粒子描述的话，则任何光束的历史亦即光子的世界线，在每一点上总沿着光锥，而任何物质粒子的历史必须在每一点的光锥的内部。这一切从图5.17可以看到。所有点处的光锥族可以被看成空间——时间的闵可夫斯基几何的一部分。

什么是闵可夫斯基几何？光锥结构是其最重要的方面。但是闵可夫斯基几何有比这更丰富的内容。它有一种和欧几里得几何的距离极相似的“距离”的概念。在三维欧几里得几何中，按照标准的笛卡儿坐标，从坐标原点到某一点的距离r可写作

$$r^2 = x^2 + y^2 + z^2 .$$

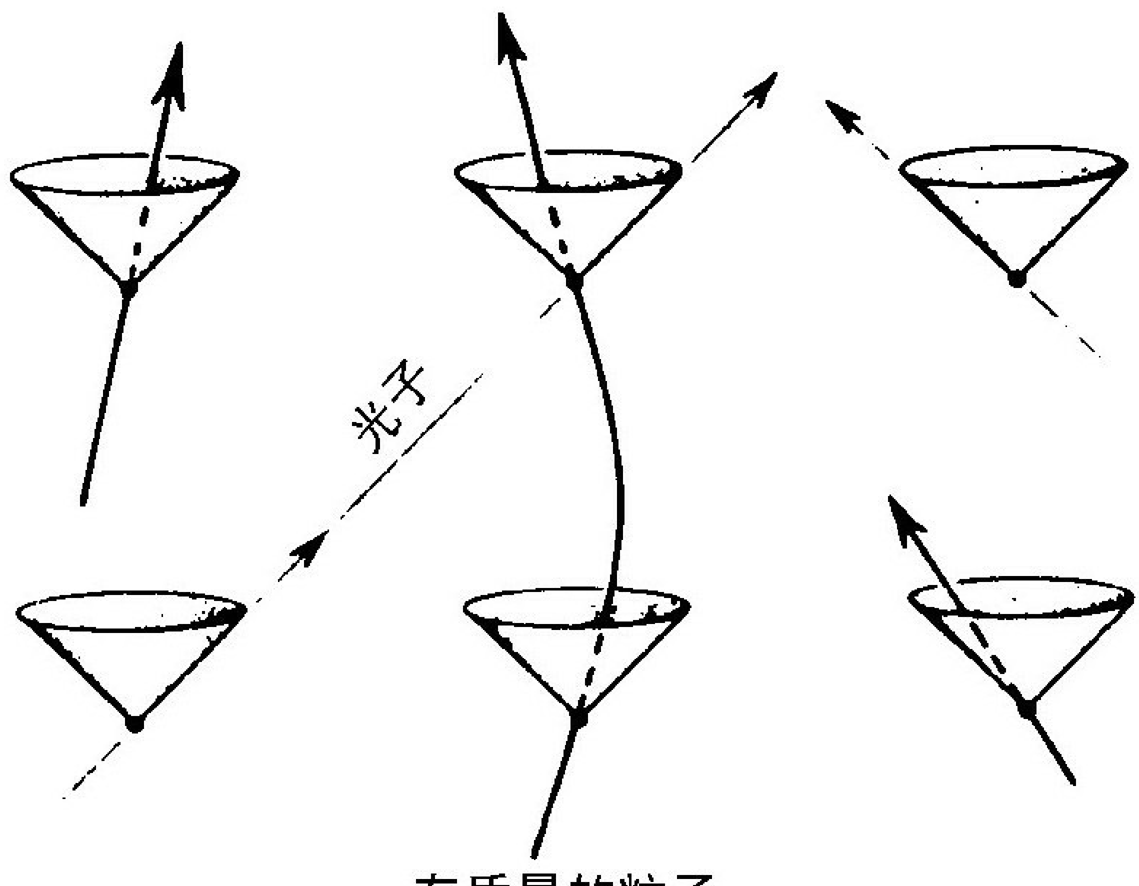


图5.17　闵可夫斯基几何图

[见图5.18（a）。这正是勾股定理——或许二维的情况更熟悉些。]在我们的三维闵可夫斯基几何中，其表达式非常相似[图5.18（b）]，根本的差别是我们有两个负号：

$$s^2 = t^2 - (x/c)^2 - (z/c)^2。$$

更正确地讲，我们应该有四维闵可夫斯基几何，当然距离表达式应写作

$$s^2 = t^2 - (x/c)^2 - (y/c)^2 - (z/c)^2。$$

此表达式中“距离”s的物理意义是什么呢？假定有一点其坐标为（t, x/c, y/c, z/c）[或者在三维的情形（t, x/c, z/c）；见图5.16]，并且在O的（未来）光锥的内部。则直线段OP可以代表某一个物质粒子——比如说由我们爆炸发射出的某一个特定粒子的一部分历史。线段OP的闵可夫斯基“长度”s有直接的物理解释，它是粒子所实际经验的事件O和P之间的时间间隔！这就是说，如果有一非常可靠和精确的钟附在该粒子上[15]，那么在事件O和事件P记录下的时间的差刚好是s。和通常预料的相反的是，坐标值t本身不描述精确的钟测量的时刻，除非它“静止”地处于我们的坐标系中（亦即x/c, y/c和z/c取固定值），这表明在图中钟有一条“垂直”的世界线。这样，只对于“静止”（亦即具有“垂直”世界线）的观察者“t”才表示“时间”。按照狭义相对论，量s为每一位从原点以均匀速度离开的观察者提供正确的时间量度。这是非常令人吃惊的——和伽利略——牛顿的简单取坐标值t为时间测量的“常识”十分矛盾。我们注意到，只要有任何运动，则相对论性（闵可夫斯基）的时间测量s总是比t要小（因为从上式我们知道，只要x/c, y/c和z/c不全为零，则s^2比t^2小）。运动（亦即OP不沿着t轴）总是使得在和坐标值t相比较的钟“变慢”。如果运动速度和c比较很小，则s和t就几乎一样，这就解释了为何我们不知道“运动着的钟走得慢”的事实。在另一种极端情况下，速度刚好为光速，P就处在光锥上，我们发现s=0。光锥刚好是它从O起，闵可夫斯基“距离”（亦即“时间”）为零的集合。这样，光子根本没有“经历”任何时间流逝！（我们不允许更极端的情况，P运动到光锥外面，因为这一来s变成虚的了——也即负数的平方根——也就是违反了物质粒子或光子不能运动得比光快的规律。[41]）

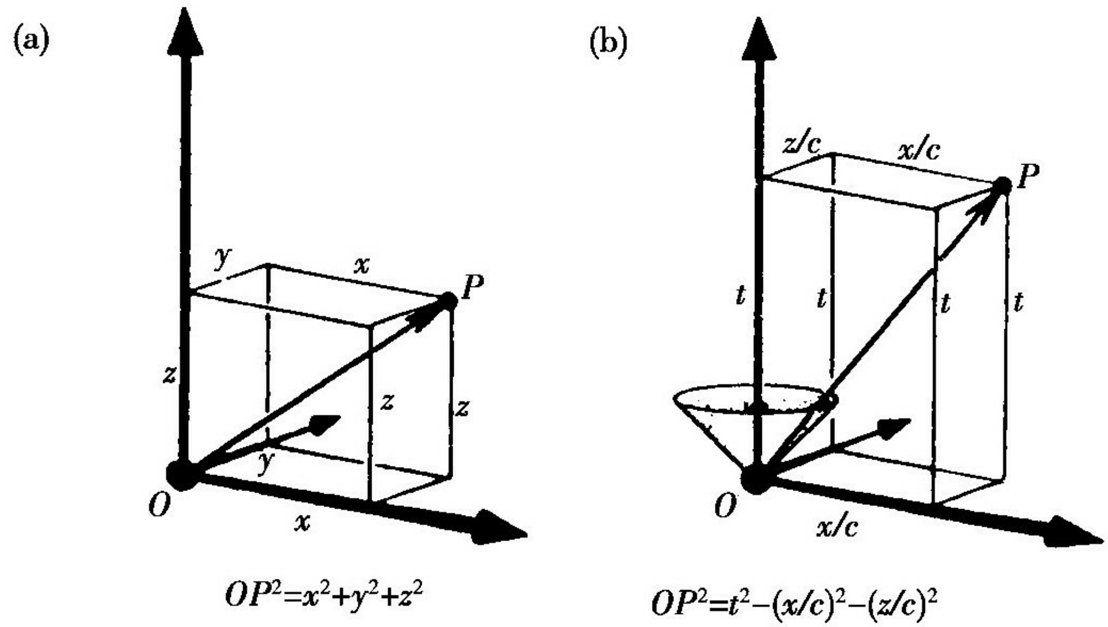


图5.18（a）欧几里得几何和（b）闵可夫斯基几何的“距离”测量的相互比较（后者的“距离”表示经历的“时间”）

可以把闵可夫斯基“距离”一样好地应用于空间——时间中的任何一对点上去，其中一点处在另一点的光锥之内——这样，一个粒子可以从一点运动到另一点。我们简单地考虑将O移到空间——时间中的某一不同点。两点间的闵可夫斯基距离是一台从一点匀速运动到另一点的钟经验的时间的间隔。当此粒子允许为光子时，闵可夫斯基距离变成零，我们两点中的一点就必须处在另一点的光锥上——这个事实可用来定义那一点的光锥。

闵可夫斯基几何的基本结构以及世界线的“长度”的古怪测度包含了狭义相对论的精华。在这里，世界线的“长度”被解释作物理钟所“测量”（或“经历”）的时间。特别是读者也许熟悉的相对论中的“双生子佯谬”：双生子中的一个留在地球上，而另一个以接近于光速的巨大速度旅行到邻近恒星上去，然后再返回。当他返回之时，人们发现两人衰老得不一样。旅行者还很年轻，而他那位待在家里的兄弟却已垂垂老矣。这按照闵可夫斯基几何很容易描述——人们可以看到，这个现象虽然令人迷惑，实际上并非荒谬。我们在图5.19中用世界线AC代表留在家中的那个双生子，而旅行者的世界线包括AB和BC两段，这代表去和回的航行的两个阶段。留在家中的那个双生子所经历的时间由闵可夫斯基距离AC所测量，而旅行者所经历的时间由两段闵可夫斯基距离AB和BC的总

和[16]给出。这两个时间不同，而且我们有

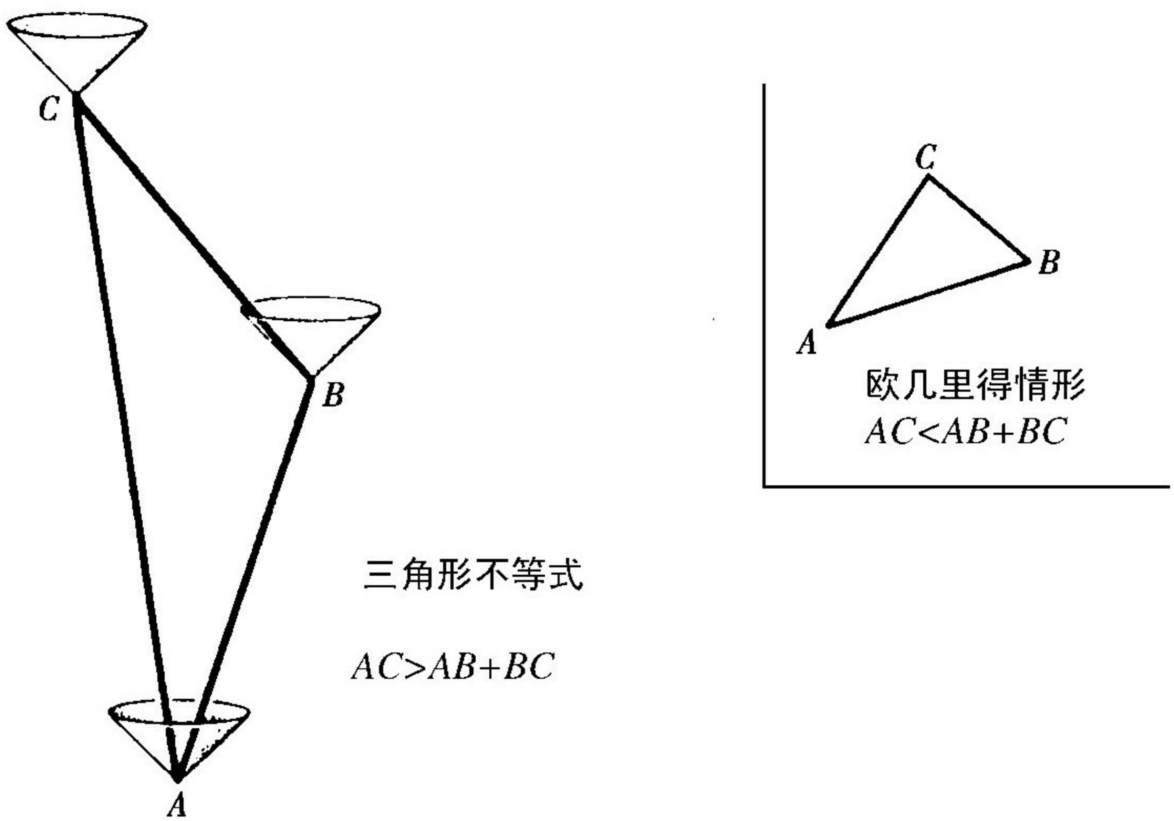


图5.19　按照闵可夫斯基三角形不等式来理解狭义相对论中所谓的“双生子佯谬”（为了比较，我们也给出了欧几里得的情形）

AC＞AB+BC，

此不等式的确表明留在家中的那个所经历的时间比旅行者更长。

上面的不等式看起来和通常的欧几里得几何中的著名的三角形不等式（A, B，C，现在变成了欧几里得空间中的三点），亦即

AC＜AB+BC

相当类似。该不等式断言，一个三角形的两边的和总比第三边大。我们并不把这个当成佯谬！从一点到另一点（这里是从A到C）之间的距离依赖于我们采取的实际途经，这是起码的常识。（在现在情形下，这两种途径为AC以及更长的折线ABC。）它是两点（此处为A和C）之

间的最短距离为连接它们的直线（直线AC）度量的特例。不等式符号在闵可夫斯基情况下的反向是因为定义“距离”时的符号改变所引起，因此闵可夫斯基的AC比折线ABC“更长”。闵可夫斯基“三角形不等式”是更一般结果的特例：连接两个事件的最长的（在经历最长时间的含义上）世界线为直线（亦即加速度为零）。如果两个双生子从同一事件A开始并终结于同一事件C。第一个双生子没有加速地从A旅行到C，而第二个加速，则他重新相遇时，前者总是经历了更长的时间流逝。

以与我们直觉相矛盾的方式，引进这样的时间测度的奇怪概念，似乎是有点荒谬。但是现在已有极大量的实验证据支持它。例如，许多亚原子粒子以一定的时间尺度衰变（亦即分裂成其他粒子）。这些粒子有时以非常接近光速的速度运动（譬如从外空间到达地球的宇宙线或是人造的粒子加速器中的粒子），它们的衰变时间精确地以从上述考虑导出的方式变迟缓。以下事实会更令人印象深刻，现代的钟（“核子钟”）可以做得如此精密，以至于时间变化效应可被快速低空飞行的飞机携带的钟直接检测出来，结果和闵可夫斯基“距离”测度s，而不和t相一致。严格地讲，考虑到飞机的高度，就牵涉到广义相对论的一个小的附加的引力效应，但是这些也都和观测相一致；（参阅下一节。）此外，还有许多其他紧密地和整个狭义相对论框架相关的效应，它们都经常接受了严密的验证。爱因斯坦的著名的关系式

$$E=mc^2,$$

即是其中之一，这表明能量和质量等效。在本章的结尾我们要遇到这一个关系式的一个令人哭笑不得的推论！

我还没有解释相对论原理如何和这类事体相协调。以闵可夫斯基几何的观点看，以不同的均匀速度运动的观察者怎么会是等同的？图5.16中的时间轴（“静止观察者”）怎么能和其他直的世界线，比如OP（“运动观察者”）完全等同？让我们先考虑欧几里得几何，很清楚，就几何整体而言，任何两条直线都是完全等同的。人们可以将整个欧几里得空间在自身上“刚性”地滑动，使得其中一条直线和另一条直线的位置重合为止。考虑一个二维亦即欧几里得平面的情形。我们可以想象在一个平面上“刚性”地移动一张纸，使得画在纸上的任一条直线和平面上的已给定的直线相重合。这个刚性运动保持几何结构不变。虽然稍不明显一些，这些议论类似地在闵可夫斯基几何中也成立。在这里人们必须小心

地理解“刚性”的含义。现在我们用一种古怪的材料取代那张滑动的纸——为了简单起见，我们首先研究二维的情况——该材料在一个45°方向上伸长而在另一个45°方向上压缩时两条45°线必须仍保持为45°线。从图5.20可看到这一点。在图5.21中我试图描绘三维的情形。这种称作庞加莱运动（或非齐次洛伦兹运动）的闵可夫斯基空间的“刚性运动”似乎显得不“刚性”，但它保持了所有的闵可夫斯基的距离。而“保持所有距离”在欧几里得情况下正是“刚性”的意义。狭义相对论原理声称，物理在这种空间——时间的庞加莱运动之下不变。尤其是，世界线为我们原先闵可夫斯基图画（图5.16）的时间轴的“静止”的观察者S和以OP为世界线的“运动”观察者M有完全一样的物理。

每一坐标平面t等于常数代表观察S的任一“时刻”的空间，亦即他认为同时（发生在“同一时刻”）的一组事件。我们称此平面为S的同时空间。当我们过渡到另一观察者M，就必须将原先的同时面族抛弃，而取代以M的同时面族 [17] 。我们注意到图5.21中的M的同时面显得向上倾斜。按照欧几里得几何的刚性运动思考，则会以为这倾斜似乎方向错了，但在闵可夫斯基情况下正是我们所预料的。当S认为所有在t为常数的平面上的事件同时发生时，M却持不同观点：从他看来，在他的每一个倾斜的等时空间上的事件才显得是同时的！闵可夫斯基几何本身并不包含“同时性”的唯一概念，而每一位匀速运动的观察者各有自己的“同时性”概念。

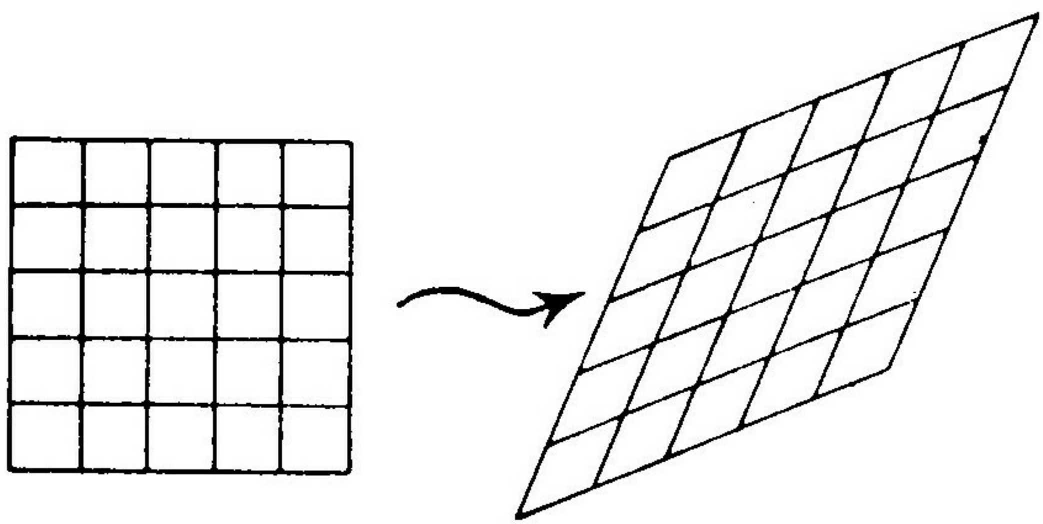


图5.20　二维空间——时间中的庞加莱运动

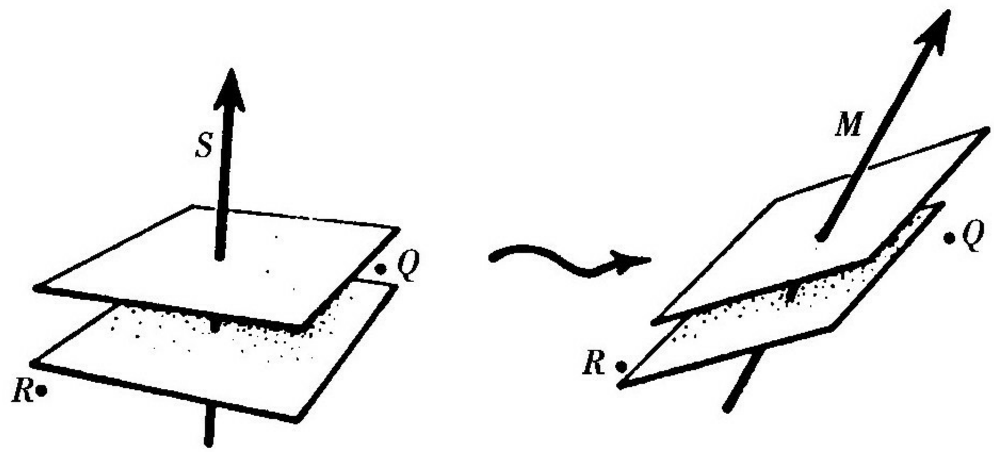


图5.21　三维空间——时间中的庞加莱运动。左图画出S的同时性空间，而右图为M的同时性空间。注意S认为R比Q早，而M认为Q比R早（此处的运动被认为是被动的，这只是因两个观察者S和M对同一空间——时间所做的不同描述所引起的）

考虑图5.21中的两个事件R和Q。依S看来，事件R在事件Q之前发生，因为R处于比Q更早的同时面上；但是，依M看来，情况刚好相反，Q处于比R更早的同时面上。这样，一个观察者认为事件R早于Q发生，而另一个观察者认为Q比R早发生！（只有当R和Q所谓类空地分隔开也就是一个事件处在另一事件的光锥之外，并因此没有物质粒子或光子能从一个事件运动到另一个事件时，这才会发生。）只要事件在相隔非常远的距离上发生，甚至非常小的相对速度也会导致重大的时序差异。假定在仙女座大星云（离开我们银河系最近的大星系，大约是2000亿亿千米那么远）处发生了一个事件，地球上两个观察者相遇时将他们的钟对好，由于他们的运动速度不同，他们俩对该事件发生时刻的判断可有几天的差别（图5.22）。对于其中一个人来说，试图去歼灭地球行星上生命的空间飞船队已上路了；而对于另外一个人来说，甚至是否要发射这个飞船队的决定都尚未做出。

爱因斯坦广义相对论

我们回忆一下伽利略关于任何物体在引力场中同样快下落的伟大的洞察。（这是洞察的而不完全是直接观察的结果。由于空气阻力作用，羽毛和石头不会一起下落！伽利略的洞察在于意识到，如果空气阻力可减少到零，它们就会一起下落。）这一直觉的深刻意义整整花了3个世纪的时间才被意识到，而成为一个伟大理论的奠基石。这就是爱因斯坦

的广义相对论——引力的一个非同寻常的描述。正如我们很快就要理解到的，为了实现它，我们需要引进弯曲的空间——时间的概念。

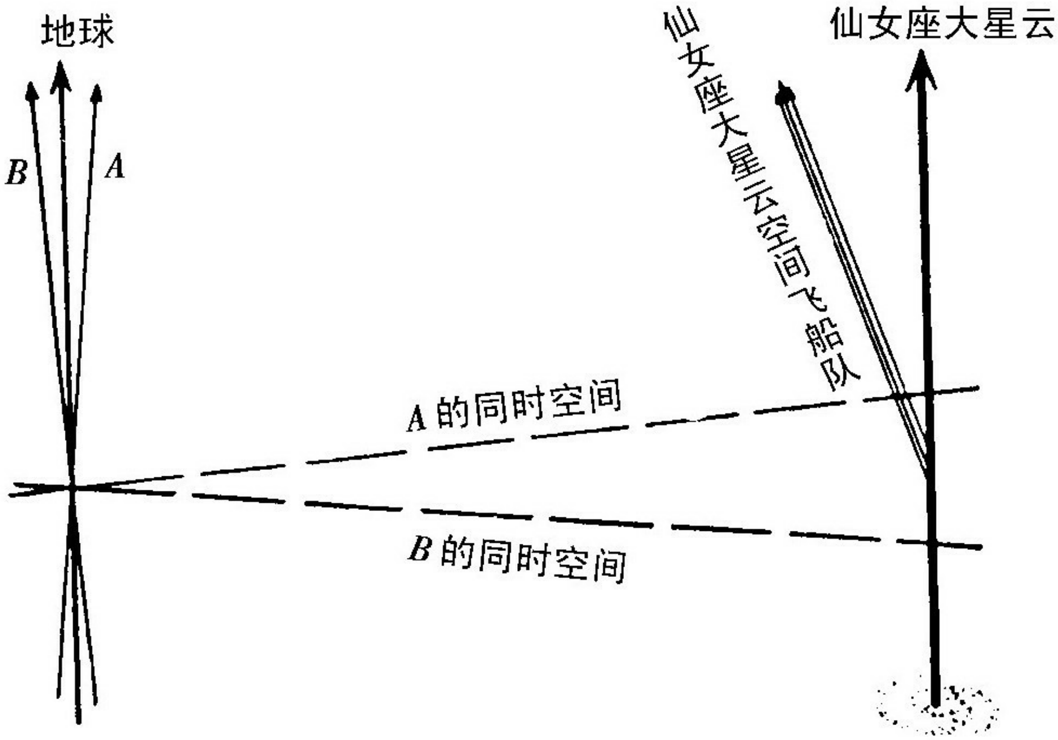


图5.22　两个人A和B相互很慢地穿过，但是他们对于仙女座大星云空间飞船队是否在他遭遇的时刻已经出发有不同的观点

伽利略的洞察和“空间——时间曲率”有何关系呢？我们知道在牛顿的理论中粒子被通常的引力所加速。这样的一个与之如此不同的思想，怎么能重新产生并且改善那个理论的所有超等的精确性呢？此外，伽利略古老的直觉包含着以后没被合并到牛顿理论中的某种东西，这怎么可能呢？

由于最后一个问题最易于回答，让我们从它开始。在牛顿理论中，是什么制约着在引力作用下的物体的加速度？首先，引力作用到物体上，牛顿引力定律告诉我们这必须和物体质量成正比。伽利略的直觉是发生在牛顿引力定律中的“质量”和牛顿定律中的是同一“质量”。（可以用“正比于”来取代“同一的”。）正是它保证了引力作用下的物体的加速度实际上与它的质量无关。在牛顿的一般理论中完全没有要求这两种质量概念的同一性。牛顿只是把它当成一个假设。的确，在平方反比律方

面电力和引力是类似的，但电力所依赖的是与牛顿第二定律中的质量完全不同的电荷。“伽利略直觉”不能应用于电力：在电场中物体（带电的物体）不会以同样的速度下落！

现在，我们就简单地接受伽利略关于引力作用下的运动的洞察，并探究其含义。设想伽利略从比萨斜塔上释放两块石头。如果在一块石头上有一镜头指向另一块石头的摄像机，那么其提供的摄像是一块在空中徘徊的石头，就像引力对它没有影响似的（图5.23）！这正是因为在重力下所有物体都以同样速度下落。

我们在这里不管空气阻力。因为在太空中实际上没有空气，所以太空飞行给我们提供了这些观念的一个更好的验证。现在，太空中“下落”简单地表示在引力作用下沿着合适的轨道运动。这个“下落”没有必要是冲着地球中心的直线下降。运动也可以有水平分量。如果此一水平分量足够大，那它就能围绕地球而不必朝向地面的方向“下落”！在引力下的自由轨道上旅行只不过是一种优雅（并且非常昂贵）的“下落”方式。正如前面使用摄像机，现在一位做“太空行走”的航天员看到他的宇宙飞船在他之前徘徊，表观上不受在他之前的地球的巨大的球体的引力的影响（图5.24）！这样，人们只要过渡到自由下落的“加速参考系”去，就可以局部地消除引力效应。

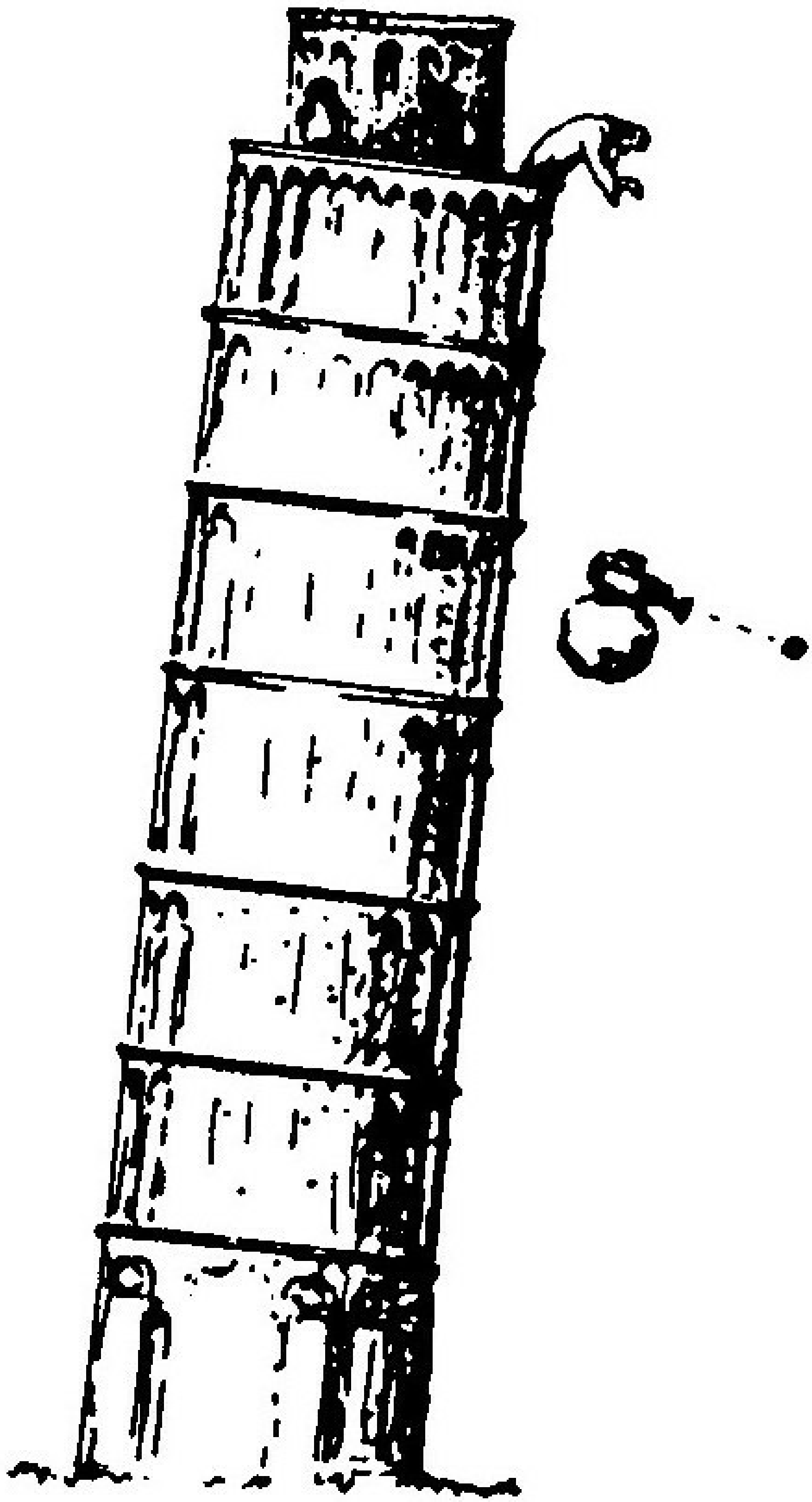


图5.23　伽利略从比萨斜塔上释放两块石头（和一台摄像机）



图5.24　航天员看到他的宇宙飞船在他之前徘徊，如同不受引力影响似的

因为引力场效应正和加速度效应一样，所以可用自由下落的方式来对消引力。事实上，你如果处在一台正在加速上升的电梯之内，就会简单地觉得表观引力场的增大；如果电梯下降，则引力场减弱。如果悬挂电梯的绳索断了，那整个下落加速度就完全抵消了引力的效应（不考虑空气阻力和摩擦效应），而电梯内的乘客就像上述的航天员那样显得在空中自由浮动，直到它撞到地面上为止！甚至在火车和飞机上，加速度会使一个人感到引力的强度和方向不和他视觉提示的应是“往下”的方向一致。这是因为加速度和引力效应是互相类似的，人的感觉不能将它们区分开来。爱因斯坦把引力的局部效应和加速度参考系的效应等效的事实称为等效原理。

上述的考虑是“局部的”。然而，如果人们允许去做足够精密的（不完全局部的）测量，他就能在原则上断定在“真正”引力场和纯粹加速度之间的区别。在图5.25中我用稍微夸张的方式显示出由许多粒子构成的原先静止的球面，在地球引力作用下自由下落时如何受（牛顿）引力场的非均匀性的影响。该引力场在两个方面不均匀。首先，因为地球在有限距离的某处，靠近地球表面的粒子向下加速比远处的粒子更快（由于牛顿平方反比律引起）。第二，由于同一个原因，在水平方向上不同位置的粒子加速度的方向也有些微差别。球面由于这种非均匀性引起了微小变形而成为一个“椭球面”。由于它靠近地球的部分遭受到比远处的部分稍微更大的加速度，它在向地球中心方向（以及相反的方向）被拉长。由于加速度在沿地球中心方向稍微向内侧的作用，它在水平方向变狭窄。

这种畸变效应被称为引力的潮汐效应。如果我们用月亮来取代地球的中心，并且粒子的球面用地球表面取代，则我们刚好得到由于月亮的影响而在地球表面产生的潮汐，鼓出的部分正是朝着和背着月亮的方向。这个不能用自由下落“消除”的引力场的一般特征正是潮汐效应。（潮汐畸变的大小实际和离开吸引中心的距离成立方反比律，而不是平方反比律的关系。）

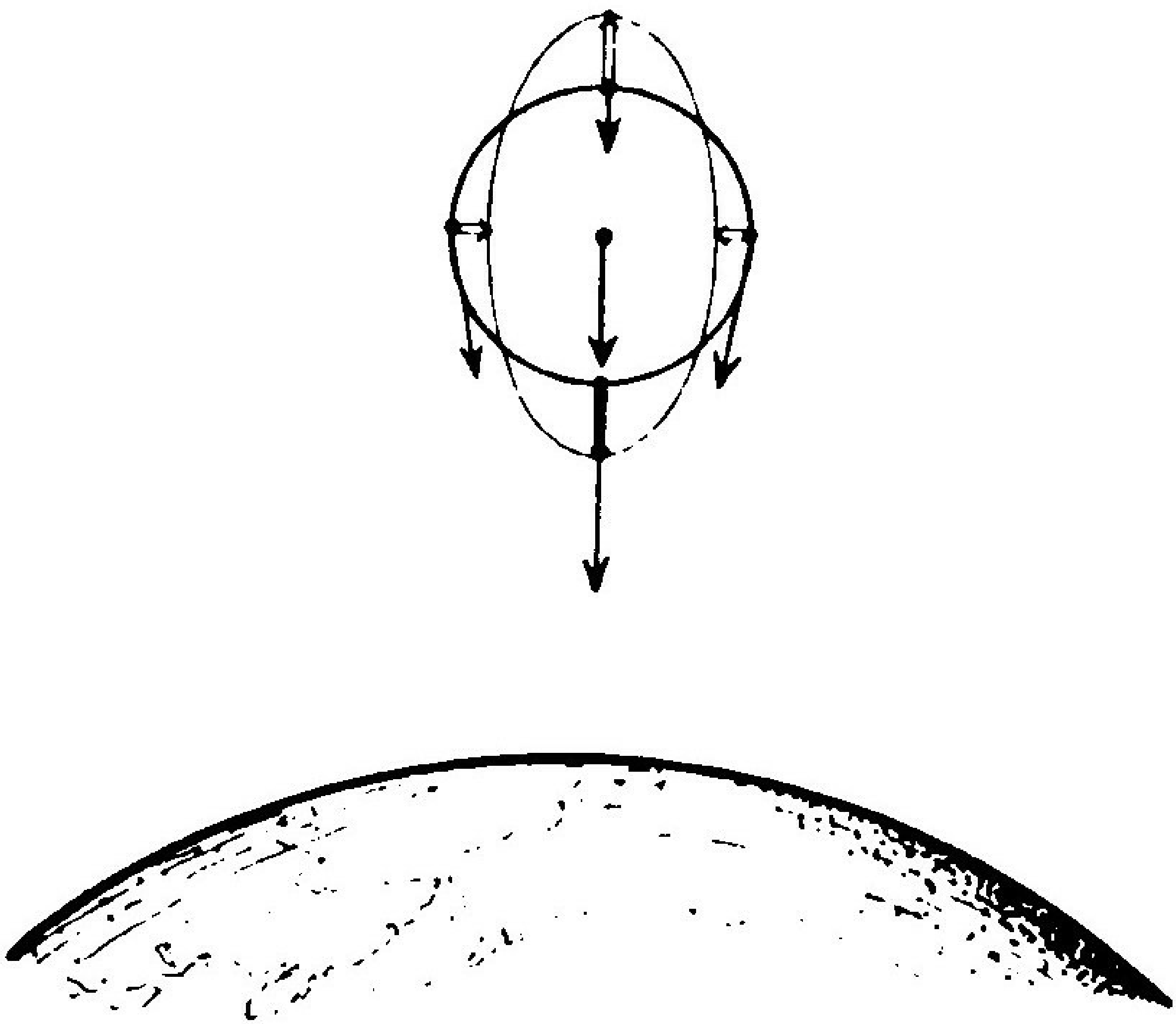


图5.25　潮汐效应。双箭头表明相对加速度（外尔）

牛顿引力的平方反比律可按照这个潮汐效应得到一个简单的解释：由原先[18]球形而畸变成的椭球的体积等于原先球体（就认为该球面围绕着真空好了）的体积。这种体积性质是平方反比律的特征，它对于其他的力的定律不成立。下一步，我们假定球面围绕着的不是真空而是总质量为M的某物体。此物体的引力产生附加的向内去的加速度分量。这样，由原先粒子球面变形成的椭球体积就会收缩，其收缩量和M成正比。我们让球面以固定的高度围绕着地球（图5.26），所发生的体积减小效应即为一个例子。由地球引力导致平常的向下（亦即向内）加速就是引起球形体积减小的同一个原因。这种体积减小效应印证了牛顿引力定律继续存在的部分，也即此力和吸引物体的质量成正比。我们画出这种情形的空间——时间图。我在图5.27上画出球面（在图5.25中画成了一个圆圈）上粒子的世界线。我在这里是用使球面的中心显得处于静止

（“自由下落”）的坐标系。广义相对论把自由下落运动看作“自由运动”——和无引力物理中的“均匀直线运动”相类似。这样，我们试图在空间——时间中用“直”的世界线来描绘自由下落。然而，从图5.27看出“直”这个字在此处的用法显得混乱。这只不过是术语的问题。我们以后就将自由下落的粒子的世界线称作空间——时间的测地线。

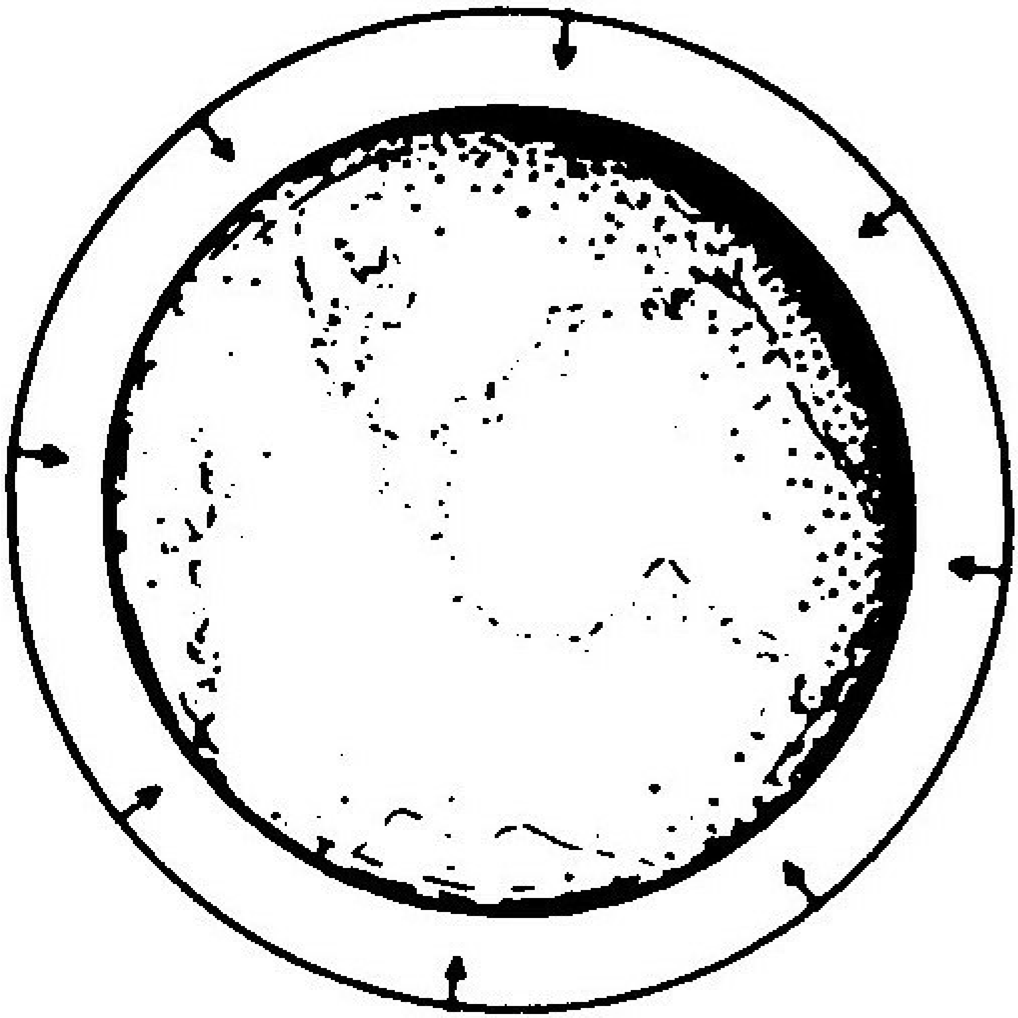


图5.26　当球面围绕着物体（此处为地球）时，就有一个纯粹向内的加速（里奇）

这是一个好术语吗？“测地线”在通常情况下的含义是什么呢？我们考察二维曲面的类似情形。测地线为在曲面上（局部的）“最短程”的曲线。如果我们想象在此曲面上拉伸一根绳子（不要太长，否则它会滑走），那么这根绳子在曲面上就和一条测地线相重合。我在图5.28上给出了两个曲面的例子，第一个具有“正曲率”（和球面类似），而第二个

具有“负曲率”（一个马鞍形的面）。在正曲率曲面上，两条互相邻近的一开始相互平行的测地线会相互靠近；对于负曲率曲面，它们会相互离开。如果我们想象，自由下落粒子的世界线在某种意义上像是曲面上的测地线，则可以看到在前面讨论的引力潮汐效应和曲面的曲率效应之间有种精确的相似性——但是现在情形下正的和负的曲率效应会同时存在。我们可从图5.25和图5.27看到，空间——时间的“测地线”在一个方向上互相离开（当它们和地球在同一直线上时）——正如图5.28中负曲率曲面的情形——在另一方向上它们互相靠近（当它们相对于地球处于水平的方向上）——正如图5.28中正曲率曲面的情形。这样，我们的空间——时间曲率确实似乎具有类似于我们两个曲面的“曲率”，但是由于更高的维数而变得更为复杂，在不同的位移上牵涉到正和负的曲率的混合。

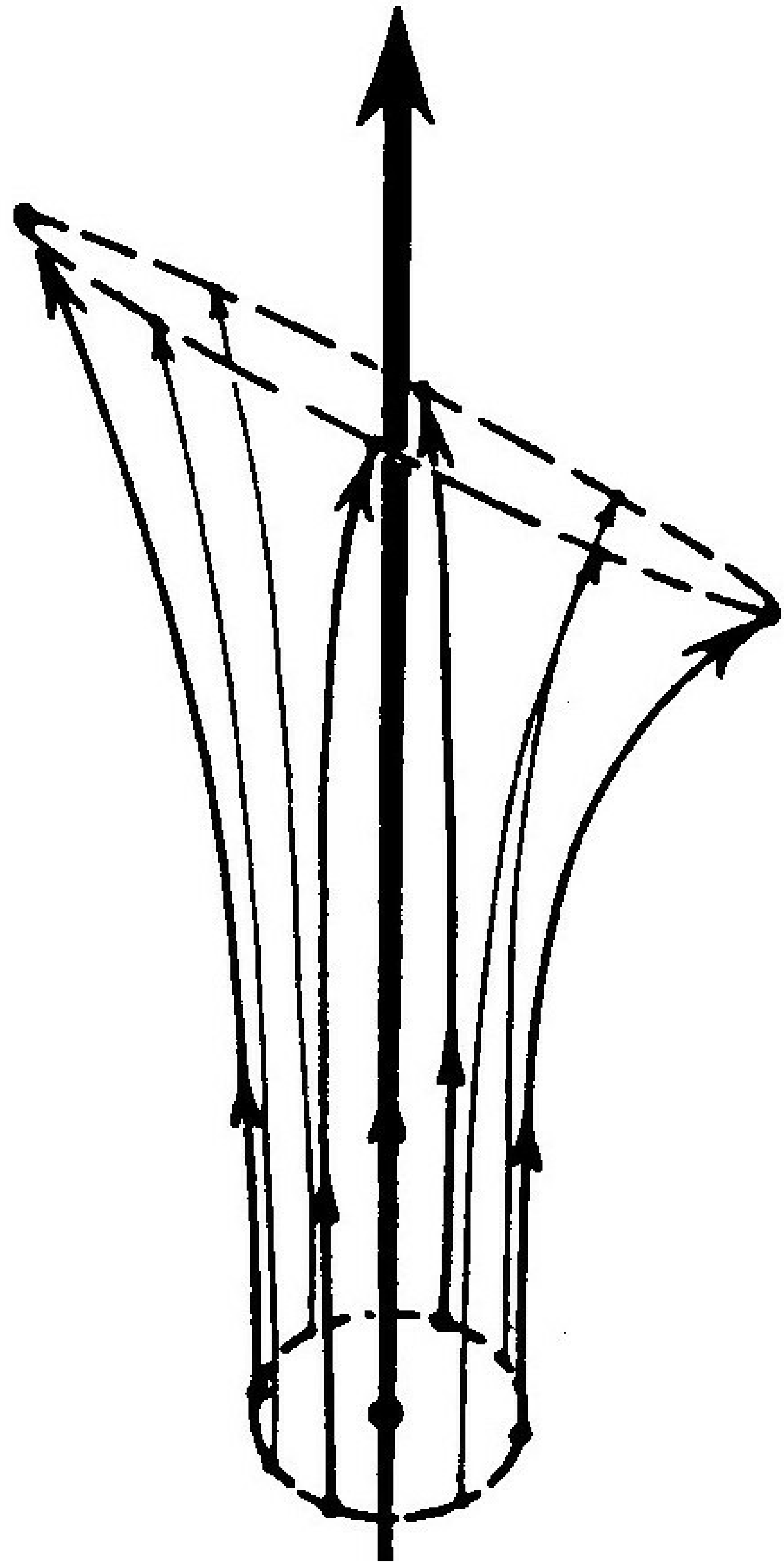


图5.27　空间——时间曲率：画在空间——时间中的潮汐效应

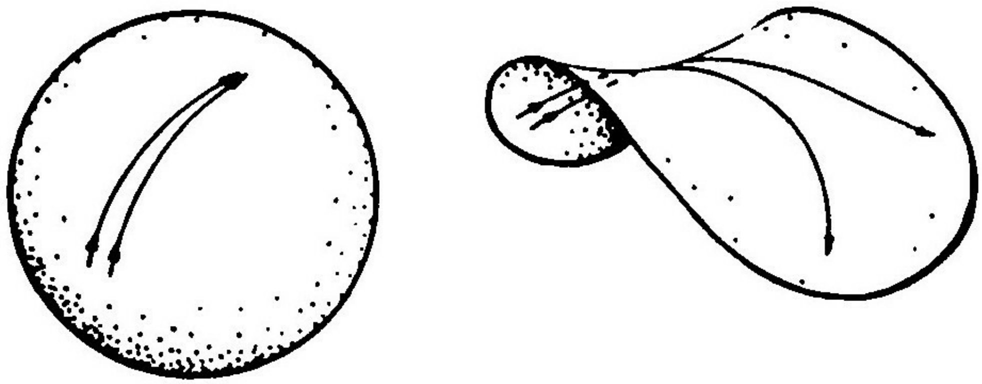


图5.28　曲面上的测地线。在正曲率处测地线收敛，而在负曲率处它们发散

这就显示了如何用空间——时间“曲率”的概念来描述引力场。这种描述的可能性归根结底是从伽利略的直觉而来的（等效原理），它允许我们用自由下落来消除“引力”。实际上我到此为止还没必要超出牛顿理论的范围。这个新的图像只是为此理论提供了重新表述 [19] 。然而，当我们将此图像和狭义相对论的闵可夫斯基描述——亦即现在我们知道应用于不存在引力情况下的空间——时间几何相结合时，就得到了新的物理。其最终的结合物即为爱因斯坦的广义相对论。

回想一下我们从闵可夫斯基得到的教益。引力不存在时，空间——时间中定义了两点之间的特殊类型的“距离”测度。我们在空间——时间中有条描述某粒子的世界线，则沿着此世界线测量的闵可夫斯基“距离”表示这个粒子实际经历的时间。（在前一节我们事实上只考虑沿着与直线段一致的世界线的“距离”，但这个断言对于任意弯曲的世界线“距离”的测量也成立。）如果没有引力场——亦即没有空间——时间曲率时，闵可夫斯基几何是准确的。但是在引力存在时，我们只能将闵可夫斯基几何当作一种近似——如同平面是弯曲曲面几何的近似描述一样。我们如果用放大倍数越来越大的显微镜去考察曲面——使得曲面的几何伸展到越来越大的范围去——则该曲面就显得越来越平坦。我们说一个弯曲曲面在局部上像是一个欧几里得平面 [20] 。我们可以以同样的方式说，在引力存在时，空间——时间在局部上像闵可夫斯基几何（也

就是平坦的空间——时间），但是我们在更大的尺度下允许某种“弯曲性”（图5.29）。特别是，正如在闵可夫斯基空间中一样，空间——时间中的任一点都是一个光锥的顶点。但是这些光锥不像在闵可夫斯基空间中的那样以完全一致的方式排列。我们将在第7章的一些空间——时间模型的例子中看到这种明显的非一致性（参阅426页的图7.13和图7.14）。物质粒子的世界线的朝向总在光锥之内，而光子的世界线总是沿着光锥。正如在闵可夫斯基空间中一样，沿着任何一条这样的曲线，总存在测量该粒子所经历的时间的闵可夫斯基“距离”的概念。正如在曲面的情形，这种距离测度定义了与平空间不同的曲面的几何。

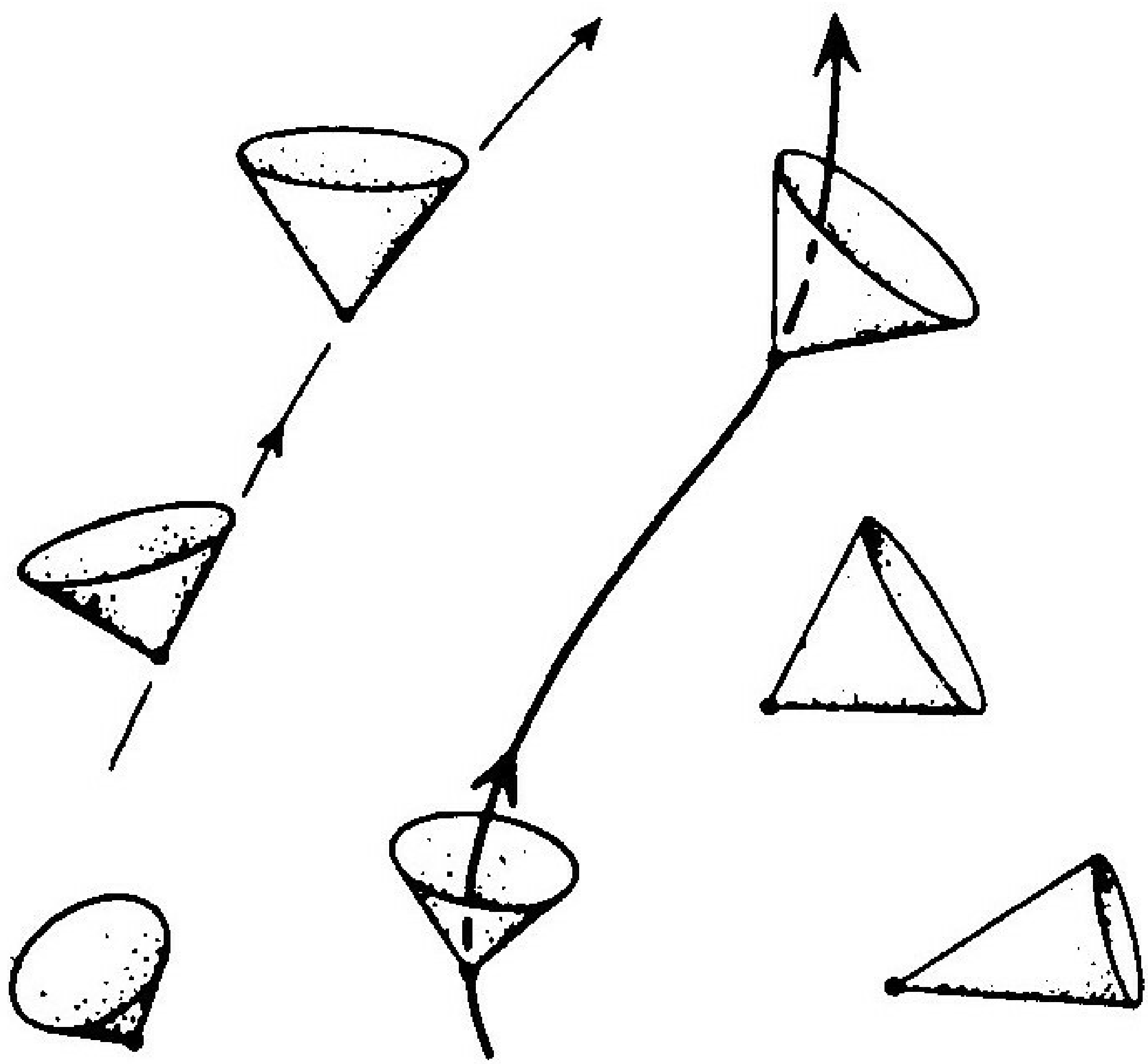


图5.29　弯曲空间——时间图

和上述的二维曲面情况相似，空间——时间中的测地线可有类似的解释。但是我们必须记住闵可夫斯基和欧几里得情形的不同之处。空间——时间中的测地的世界线取（局部）最大的距离（亦即时间），而不是取（局部）最小的长度。按照这一规则，引力作用下的自由运动粒子的世界线，实际上是测地线。这样，尤其是在引力场中运动的天体可用测地线来描写。在空虚的空间的光线（光子的世界线）也是测地线，并且是具有零“长度”的测地线[21]。我在图5.30中作为例子画出了地球和太阳的世界线的略图，地球绕太阳的运动是一根绕着太阳世界线的螺旋状的测地线。我也标出了从一个遥远的恒星到达地球的光子。因为按照爱因斯坦理论，光线被太阳的引力场所偏折，所以其世界线显得稍微有些“弯折”了。

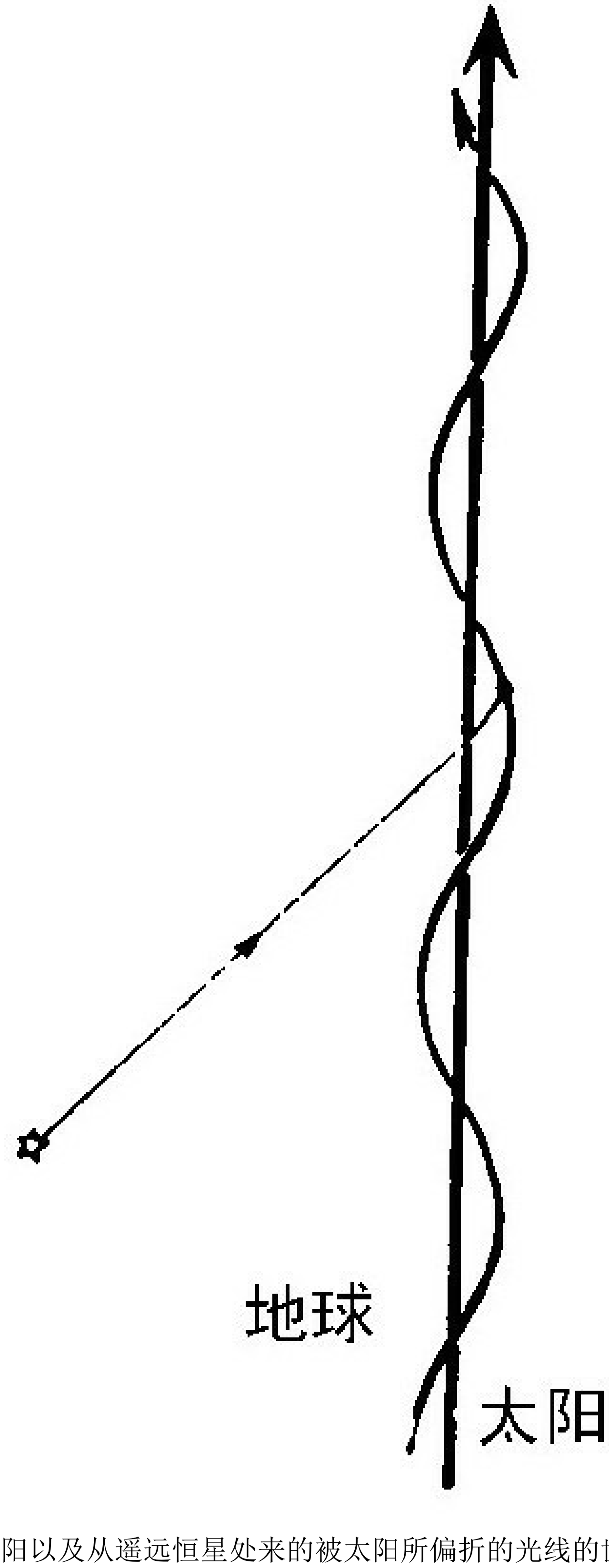


图5.30　地球和太阳以及从遥远恒星处来的被太阳所偏折的光线的世界线

我们还要看看如何将牛顿的平方反比律包括进来，并按照爱因斯坦相对论作何种修正。让我们回到在引力场中下落的粒子球面的例子上来。我们记得，如果球面围绕的只是真空，则按照牛顿理论，球的体积一开始不会改变；但是如果围绕的是一个总质量为M的物体，则会产生和M成正比的体积减小。这种规律在爱因斯坦理论中（对于小球面）刚好是一样，除了决定此体积减小的不完全是M，还有一附加的“通常非常小的”来自被围绕物质的压力的贡献。

四维空间——时间的曲率必须描写在任何地方在任何可能方向运动的粒子的潮汐效应。它的完整数学表达式由被称为黎曼曲率张量的量所给出。这个东西是有点复杂，在每一点具有20个称作分量的数。不同的分量是在空间——时间中在不同方向上的不同曲率。黎曼曲率通常写作 R_{ijkl} 。但是因为我不想在这里解释这些小指标的意义（事实上也不想解释张量的意义），我就简单地将它写作：

<u>黎曼。</u>

存在一种将此张量分解成两部分的方法，第一部分是外尔张量，第二部分是里奇张量（各有10个分量）。此分解可表达如下：

<u>黎曼=外尔+里奇。</u>

（其具体表达式在目前并不特别有用。）外尔张量外尔是测量我们自由下落的球面的潮汐畸变（亦即形状的初始变形，而非尺度的变化），而里奇张量里奇测量其初始体积改变 [22] 。我们记得，牛顿引力理论要求下落球面所围绕的质量和这初始体积的减小成正比。粗略地讲，它告诉我们，物体的质量密度，或等效的能量密度（因为 $E=mc^2$ ）——应该和里奇张量相等。

事实上，这基本上就是广义相对论的场方程——也即爱因斯坦场方程——实际的断言 [23] 。然而，关于这些还有许多技术上的细节，最好不在这里纠缠。只要知道存在一个称作能量——动量的张量，它将有关的物质和电磁场的能量、压力和动量都组织在一起。我把这一张量叫作能量，则爱因斯坦方程可非常粗略地写作：

里奇=能量。

正是在能量张量中“压力”的出现以及为了使整个方程协调的条件要求，使得压力正如前述的也对体积缩小效应有所贡献。

此方程似乎没有牵涉到外尔张量。但它是一个重要的量。在空虚的空间里感受到的潮汐效应纯粹是由外尔引起的。事实上，上述的爱因斯坦方程意味着，存在将外尔和能量相联系的微分方程，颇像我们以前遇到的麦克斯韦方程[24]。的确，把外尔当作用E、B这一对量描述的电磁场量（实际上也是一个张量——麦克斯韦张量）的引力类似物是一种富有成果的观点。在一定的意义上可以讲，外尔实际上是引力场的测定。外尔的“源”是能量张量。这和电磁场（E, B）的源是（ρ，j），也即麦克斯韦理论的电荷和电流的组合的情形很相似。这种观点将有助于第7章的讨论。

如果注意到在爱因斯坦理论和牛顿在两个半世纪前提出的理论之间，虽然在形式和内在的观念之间有如此深刻的差别，但在观测上要找到差异却非常困难，人们会十分惊异。假如所考虑的速度和光速c相比较小很多，并且引力场不太强（使得脱逸速度比c小得多，参阅第7章422页），那么爱因斯坦理论的结果实质上和牛顿的一样。但是，在这两个理论的预言的确不同时，爱因斯坦理论更准确。现在已有几个令人难忘的实验，证明爱因斯坦新理论完全成立。正如爱因斯坦所坚持的，在引力场中钟走得慢一些，此效应以不同的方式得到直接的测量。光和无线电波的确被太阳所偏折，并被遭遇者稍微地延迟——也很好地检验了广义相对论效应。空间探测器和运动行星，正如爱因斯坦理论所要求的那样，对牛顿轨道要做小修正，这些也被实验所证实。（特别是从1859年起天文学家就开始忧虑的被称作“近日点进动”的水星运动的失常，1915年为爱因斯坦所解释。）也许最令人印象深刻的是，对一个包括一对微小的大质量恒星（假定为两个“中子星”，参阅422页）的称作脉冲双星系统上的一系列观测，其数据和爱因斯坦理论非常接近，并间接地证实了一个在牛顿理论中根本不存在的效应，即引力波的辐射。（引力波是电磁波的引力类似物，以光速c来传播。）还没有找到任何被确证的和爱因斯坦广义相对论相冲突的观测。正因为这种种奇异的现象，使我们坚信爱因斯坦理论是对的！

相对论因果性和决定论

我们记得在相对论中，物质不能运动得比光快——也就是说，它们的世界线必须处于光锥之中（图5.29）。（尤其在广义相对论中，我们必须用这种局部的方式描述事物。光锥并不均匀地排列着，所以讲非常远的粒子的速度是否超过这里的光速并没有多大意义。）光子的世界线沿着光锥，但对于任何粒子都不能允许其世界线处在光锥之外。事实上，更一般的陈述应是，不允许任何信号在光锥外传播。

要理解为什么这样，可以参考闵可夫斯基空间图（图5.31）。假定我们有一台能发出比光传播得更快的信号的仪器。利用这台仪器，观察者W从他的世界线的事件A发出一个到达遥远的事件B的信号，B刚好处于A的光锥的下面。从W的观点看，可以画成图5.31（a）的样子。但从第二个观察者U的观点看，应重新画成图5.31（b）的样子，U正在进行着离开W（譬如讲，从AB之间的某点开始）的快速运动。对于U而言，事件B显得比A还更早地发生！[正如前面（257页）提到的，这种“重画”是一个庞加莱运动。]从W的观点看，U的同时性空间看起来是“向上倾斜”的，这就是为何事件B从U的观点看显得比A还早的原因。这样，对于U而言，W似乎是在往时间向后的方向上发出信号！

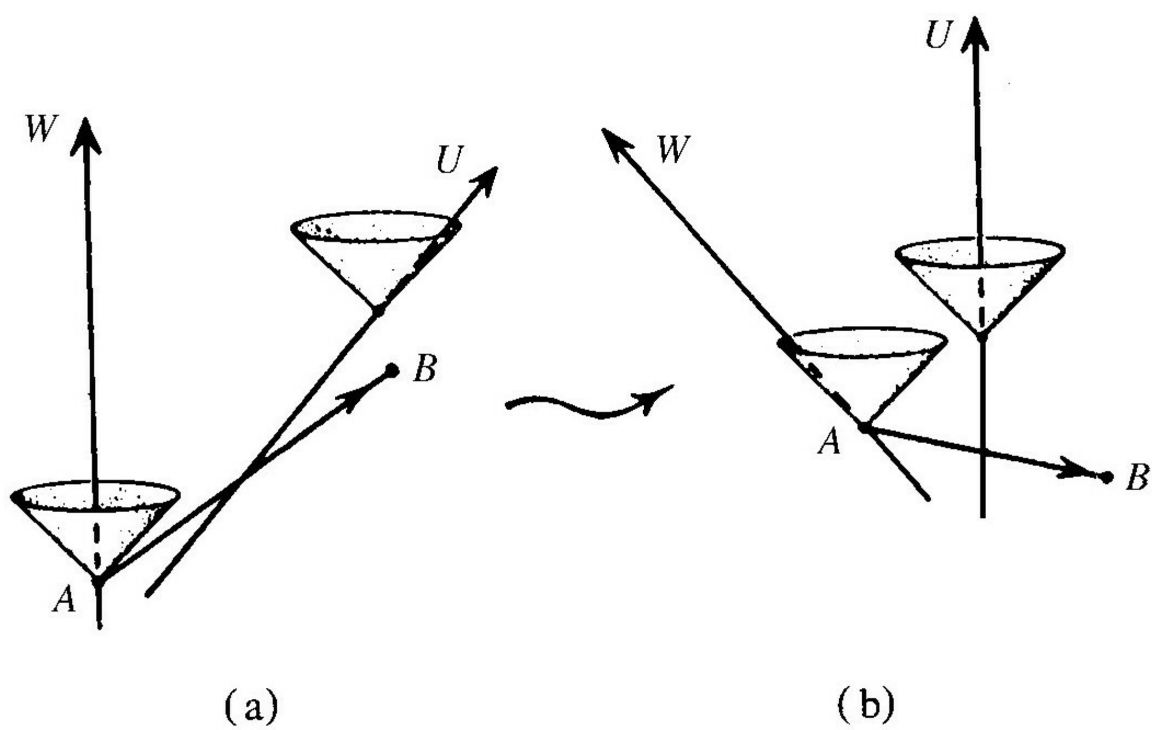


图5.31　从观察者W看来比光更快的信号，在观察者U看来变成在时间上向后行进。右图（b）只不过是左图（a）以U的观点重新画出（这种重画可视作庞加莱运动。可将其和图5.21相比较——但这里以（a）到（b）是采用积极的而非消极的意义上的变换）

这还不算什么矛盾。但是如果还有一个从U的观点看对称的（由于狭义相对论原理），离开U以和W相反的方向运动并装备有与W一样的仪器的第三个观察者V，他也能发出一个刚好比光还快的信号。从他（亦即V）的观点看，该信号是向U的方向返回。从U的观点看来，这信号又是发向过去，但这回是沿着相反的空间方向。V可以在接到W发出的原始信号的B时刻发出第二个信号到W去。从U看来，该信号在比原先发射事件A更早的事件C处到达W（图5.32）。但比这更糟糕的是，实际上事件C在W自身的世界线上比事件A更早，W在发出A信号之前即经历了事件C！观察者V发回到W的信号由于W的预先安排，可以简单地重复B处收到的。这样，W就会在自己的世界线更早的时刻收到后来想发出的同一个信号！将两个观察者分隔足够大的距离，我们就可以使得返回信号比原始的信号早一个任意长的时间间隔。也许W原始的信号是说他折断了腿，他可在此事件发生之前接受到返回信号，然后（假定）用他自己的意志，采取行动去避免事故发生！

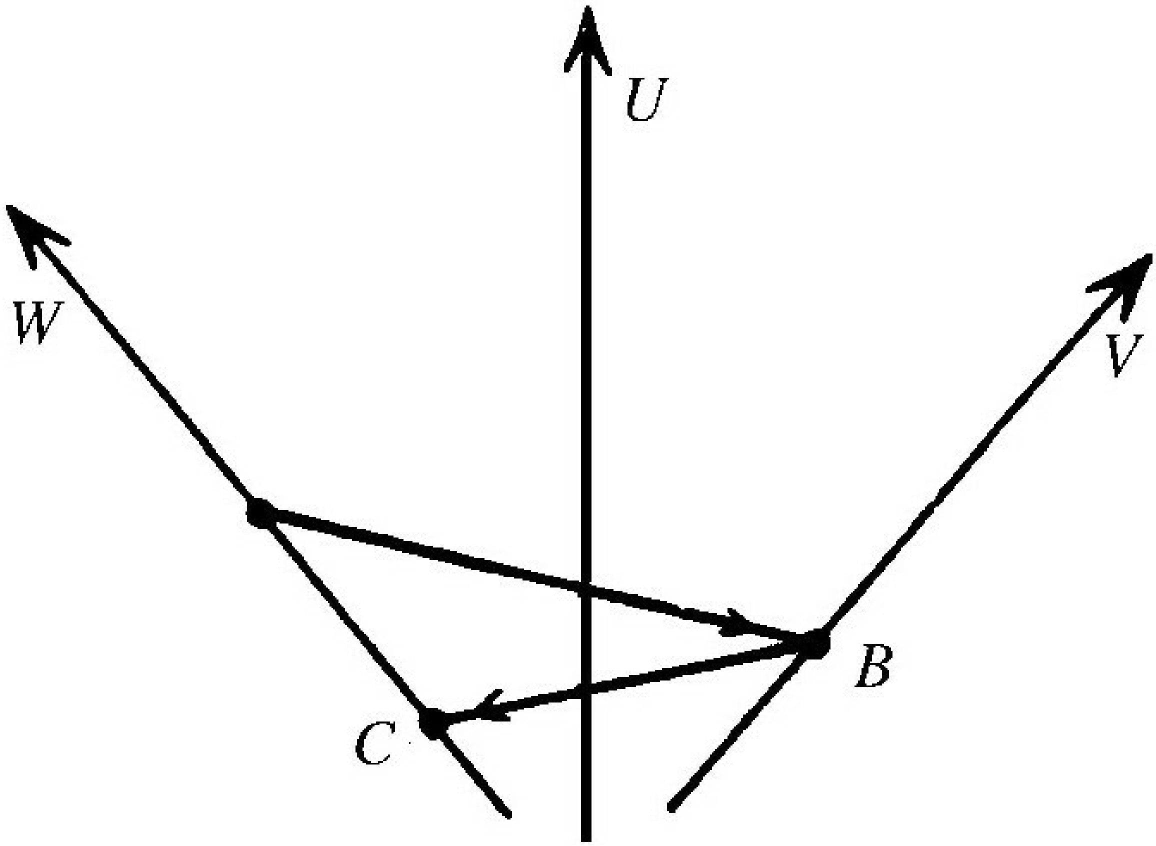


图5.32　如果V装备有和W一样的超光速信号的仪器，但该仪器的指向和W的相反，它就可以被W用来向他自己的过去发送信息

这样，先知先觉地发射信号和爱因斯坦的相对论原理一道会导致和我们“自由意志”的正常感觉的严重冲突。实际上的情形比这还要更严重。因为我们可以设想，也许“观察者W”仅仅是一台机械仪器，它的程序是如果收到“不”的信号时即发出“是”的信号，反之亦然，而V也可以是一台机械仪器，如果收到“不”的信号时即发出“不”的信号，反之亦然。这就导致了和我们以前遇到的 [25] 同样的矛盾。现在似乎和观察者W是否有自由意志“无关”，并且告诉我们超光速信号发射仪器不存在物理学上的可能性。这会在下面给我们带来一些令人困惑的推论（第6章366页）。

让我们接受，任何种类的信号——不仅仅是通常物理粒子所携带的——必须被光锥所限制。上面的论证实际上只牵涉到狭义相对论。但是在广义相对论中，这一个狭义相对论的规则仍然定域地成立。正是狭义相对论的这种局部有效性告诉我们信号必须被光锥所限制，所以它也应该适用于广义相对论。我们将会看到这一点如何影响这些理论决定论的

问题。我记得在牛顿（或哈密顿等）理论中，“决定论”意思是说在一特定时刻的初始值完全固定了其他时刻的行为。如果在牛顿理论中采用空间——时间的观点，则给定初始值的那个“特定时刻”即是四维空间——时间中的某一个三维“截面”（亦即那一时刻的整个空间）。在相对论中，不可能为此而挑出一个全局的“时间”概念。通常的步骤是采用一种更灵活的做法。任何人的“时间”都可以。在狭义相对论中，可采取某个观察者的同时面，并用此同时面来取代上述的“截面”以赋予初始值。但在广义相对论中，“同时空间”的概念并没有很好地定义。从而人们使用更普遍的类空面 [26] 的概念。我们在图5.33画出了这样的一个面；它的特征是处于它上面的每一点的光锥之外——这样，在局部上它和同时空间很相似。

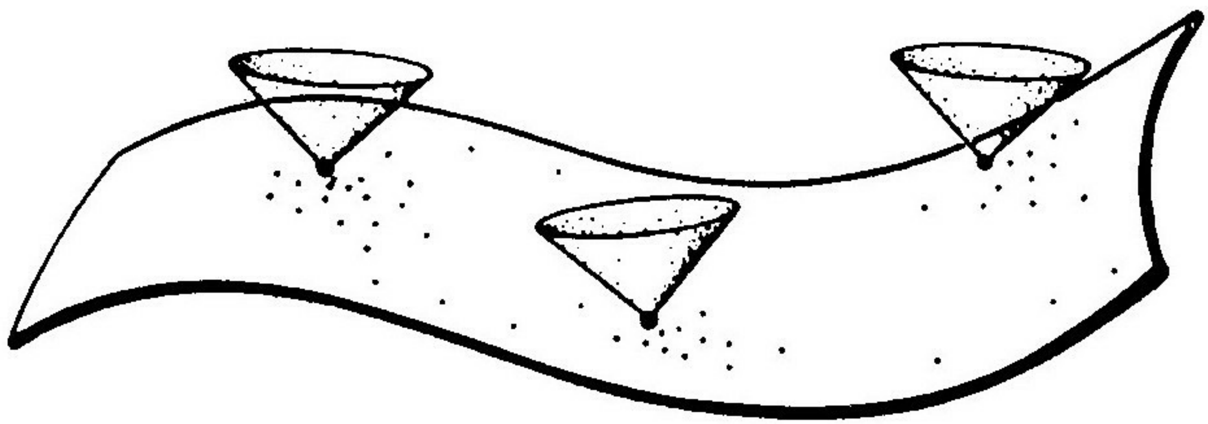


图5.33　在广义相对论中被挑选来赋予初始值的一个类空面

在狭义相对论中，决定论可以表述成为在任何给定的同时面S上的初始值，固定了整个空间——时间中的系统的行为的这一事实。（尤其是在麦克斯韦理论中这一点成立——它的确是“狭义相对性”的理论。）然而，人们可以有更强的陈述。如果想知道处在S的未来的某一事件P处发生的事，则只需要知道S上某一（有限的）有界的区域内，而不必是整个S上的初始值即可。这是因为“信息”不能传递得比光还快，而S上的任何离得太远的以至于光信号不能到达P的点不能对P有何影响（图5.34） [42] 。这实际上比在牛顿理论中出现的情形更令人满意。在那里，人们为了能对将来某一时刻要发生的事件做任何预言，原则上要知道整个无限的“截面”上发生的事。牛顿式信息的传播速度不受任何限制，牛顿的力是瞬息性的。

广义相对论中的“决定论”比在狭义相对论中复杂得多，我在此只做少许评论。首先，我们为了赋予初始值必须使用一个类空面S（不仅仅是一个同时面）。人们发现，如果像通常那样假定对能量张量有贡献的物质场的行为是决定性的，则爱因斯坦方程的确给出了引力场的局部的决定性的行为，然而，这里事情相当复杂。空间——时间的几何自身——包括它的光锥的“因果性”结构——现在成为实际上要被确定的一部分。由于我们预先不知道光锥结构，所以不能得知S的那一部分为确定未来某一事件P的行为所必须。在某种极端的情况下，甚至有整个S都不够的情形，而因此就损失了全局的决定性！这里牵涉到非常困难的问题，它们和一个在广义相对论中称为“宇宙监督”的未被证明的猜测相关。这猜测和黑洞形成有关系（参阅提普勒等1980年；参阅第7章423页以及425页处的脚注和436页）。情况似乎很可能是，和“极端”的引力场的情形相共存的“决定性失效”和人类尺度的事件几乎没有任何直接关系。但是，从这里也可以看出，广义相对论中的决定论的问题绝不像人们设想的那样干脆利落。

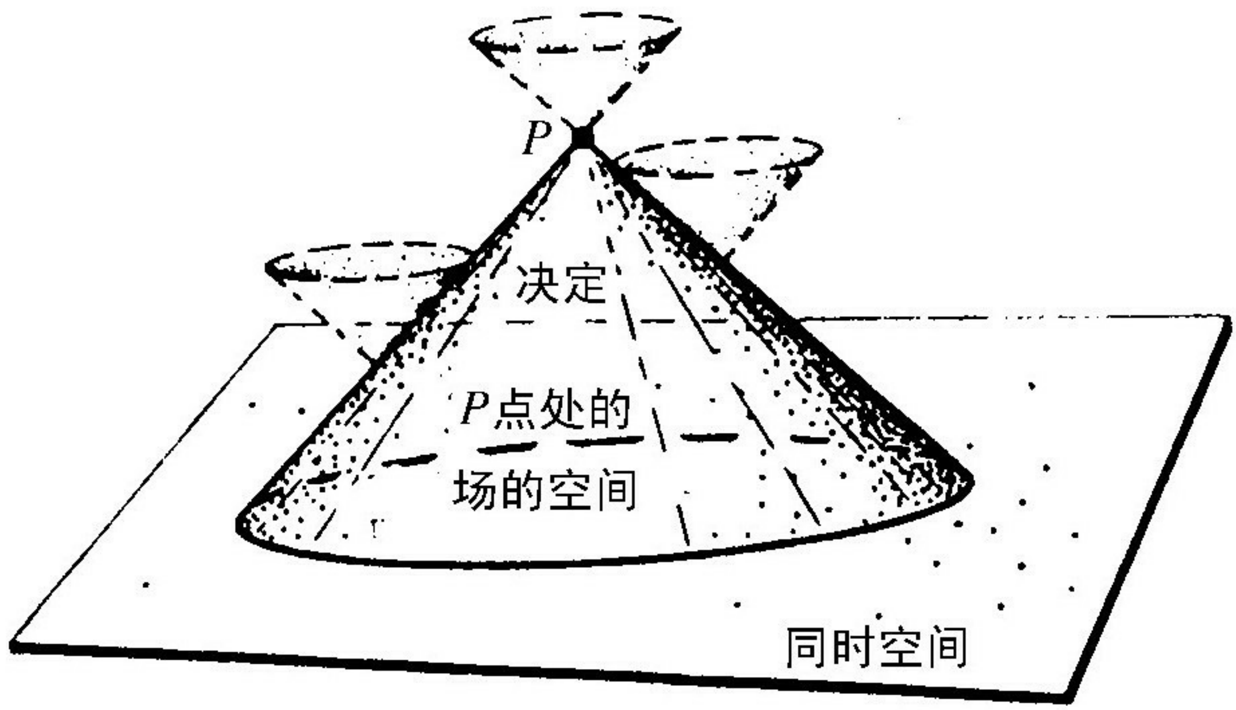


图5.34　在狭义相对论中，发生在P的事件只依赖于在同时空间中的一个有限区域的数据。这是因为传递到P的效应不能比光走得更快

经典物理的可计算性：我们的立场如何

我在这一章从头到尾，不但总是要同时留心与决定论不同的可计算性的问题。而且，我还要试图指出，在谈论到“自由意志”和精神现象时，可计算性的问题至少和决定论性的问题一样重要。但是，正如我们不得不相信的那样，在经典理论中决定论本身也不是那么清楚的。我们看到了带电粒子，运动的经典洛伦兹方程所引起的一些困扰的问题。（回忆狄拉克的“脱逸解”。）我们还注意到，在广义相对论中存在一些决定论的困难。在这些理论中，只要没有决定论，当然也就不可计算了。然而上面引用的情形中似乎没有一种因为缺乏决定性而和我们有许多直接的哲学方面的关系。在这些现象中还是没给我们的“自由意志”留下余地：在第一种情况，因为点电荷的经典洛伦兹方程（正如狄拉克解决的那样）被认为在提这些问题的水平上在物理上不合理；第二种情况，由于经典广义相对论所引起的这些问题（黑洞等）的尺度和我们自己大脑的尺度差别太大。

现在，我们在经典理论中关于可计算性的境况如何呢？可以合理地猜测，如果超越了我刚才提出的因果性和决定性的差别的话，则广义相对论中的情形和狭义相对论不会有大的差别。任何在物理系统的未来行为被初始值所决定的地方，用我们在牛顿理论情况下类似的推论，则其未来的行为似应也被那些数据可计算地决定[27]（除了上面考虑过的，玻——埃勒——里查德遭遇到的波动方程的不可计算性的“无益的”非可计算性的类型——这种情况对于光滑地变化的数据不会发生）。的确，在我迄今讨论过的任何物理理论中，很难看到任何重大的“不可计算”的因素。可以肯定预料到的是，在这许多理论中会发生“混沌的”行为，只要初始数据做非常微小的改变，就会对结果的行为产生巨大的影响。（看来在广义相对论中真是如此，参阅*Misner 1969*，*Belinskii et al.1970*。）但是，正如我在前面所提到的，很难看出这类不可计算性亦即“不可预言性”对要“驾驭”物理定律的可能的不可计算因素的仪器有何“用处”。如果“大脑”可以任何方式利用不可计算的因素，那么这种因素必须是非经典物理的。我们需要在浏览了量子理论之后，重新回来审查这个问题。

质量、物质和实在

让我们简略地清查一下经典物理所呈现的世界图像。首先空间——时间担负着主要任务：提供舞台给所有不同的物理现象。其次是任意不

停活动着的物理对象，但这些活动由精密的物理定律所约束。共有两类物理对象：粒子和场。关于粒子，除了各个都有自己的世界线以及具有各自的（静）质量和也许还有电荷等，我们很少提到它们的实际性质或特殊品质。另一方面，场的特性非常明确——服从麦克斯韦方程的电磁场以及服从爱因斯坦方程的引力场。

在处理粒子时存在一种互相冲突的情形。如果粒子的质量是如此微小，以至于其对场的影响可以忽略，则可称作检验粒子——而它们对场的响应的运动是毫不含糊的。洛伦兹力定律描述检验粒子对电磁场的响应，而测地线定律描述它们对引力场的响应（如果两种场都存在时，是上述情形的适当的结合）。这些粒子在这里必须被认为是点粒子，也就是具有一维的世界线。然而，当粒子对场（并因而对其他粒子）的效应必须考虑时——亦即，这些粒子成为场的源时——那么该粒子必须认为是在某种程度上在空间中散开的对象。否则在每个粒子的紧邻处的场会变得无穷大。这些散开的源为麦克斯韦方程提供了所需要的电荷——电流分布（ρ，j），也为爱因斯坦方程提供了所需要张量能量。除此之外，所有这些粒子和场所处的空间——时间具有直接描绘引力的可变的结构。“舞台”参与到在它上面表演的情节中去！

这就是经典物理在有关物理实在的性质方面给我们的教导。很清楚，我们在中学学到了许多，但同时我们又不可过于自得，以为我们一时形成的图像不会被某种以后更深刻的观点所推翻。我们在下一章会看到，甚至相对论所带来的革命性变革在与量子力学相比较时都会显得黯淡无光。但是，我们和经典理论以及它对物质实在的描述方面缘分还未尽。还有件使我们惊奇的事！

什么是“物质”？它是实际的物理对象，亦即世界的“东西”由之构成的实体。它是你、我以及我们的房子由之所组成的材料。如何量化物质？初等物理教科书为我们提供了牛顿的清楚的答案。它是一个对象或一群对象的质量，它是所包含的物质的测度。这看来的确是对的——没有任何其他的物理量能在作为总物质的真正量度这一点上和质量认真地作较量。况且它是守恒的：任何系统的质量，也就是物体内容的总量总是保持不变。

爱因斯坦狭义相对论中的著名公式

$$E=mc^2$$

还告诉我们质量（m）和能量（E）是可以互换的。例如，一个铀原子会衰变分裂成小块，如果能够使这些小块处于静止，则这些小块的总质量会比原来铀原子的质量小；但是若把每一块的运动的能量——动能（参阅214页[43]）——也计算在内，再除以c^2（因为$E=mc^2$）以转化为质量值，则我们发现总量实际上是不变的。质量的确是守恒的，但由于部分是由能量组成，它作为实在物质的量度显得不那么清楚了。能量毕竟依赖于物质运动的速度。一列直达列车的运动的能量相当大，但是如果我们刚好坐在此火车上，则按照我们自己的观点，火车根本没有运动。运动的能量（虽然单独粒子的杂乱运动的热能不会）会因为适当地选择观点而被“减少到零”。一种称作π^0介子的亚原子粒子的衰变是一个鲜明的例子，爱因斯坦的质量——能量关系的效应在这个场合达到了极致的程度。它肯定是一种具有定义得很好的（正的）质量的物质粒子。大约10^{-16}秒之后，它几乎总是分解（像上述的铀原子那样，但要更快速得多）成仅仅两个光子（图5.35）。从和π^0介子一起处于静止的观察者看来，每个光子携带走一半能量，这的确是π^0介子质量的一半。然而这光子“质量”具有一些模糊的性质：它是纯能量。如果我们能在一个光子的方向上快速地运动，我们就能将其质量——能量要减小到什么程度就减小到什么程度——光子的内禀质量（或正如我们很快就要讲到的静质量）实际上为零。所有这一切为质量守恒描绘出一幅协调的图像，但是它和我们过去的不完全一样。在某种意义上，质量仍是“物质的量”的测度，但在观点上有显著的改变：既然质量等效于能量，那么系统的质量，正如能量那样依赖于观察者的运动！

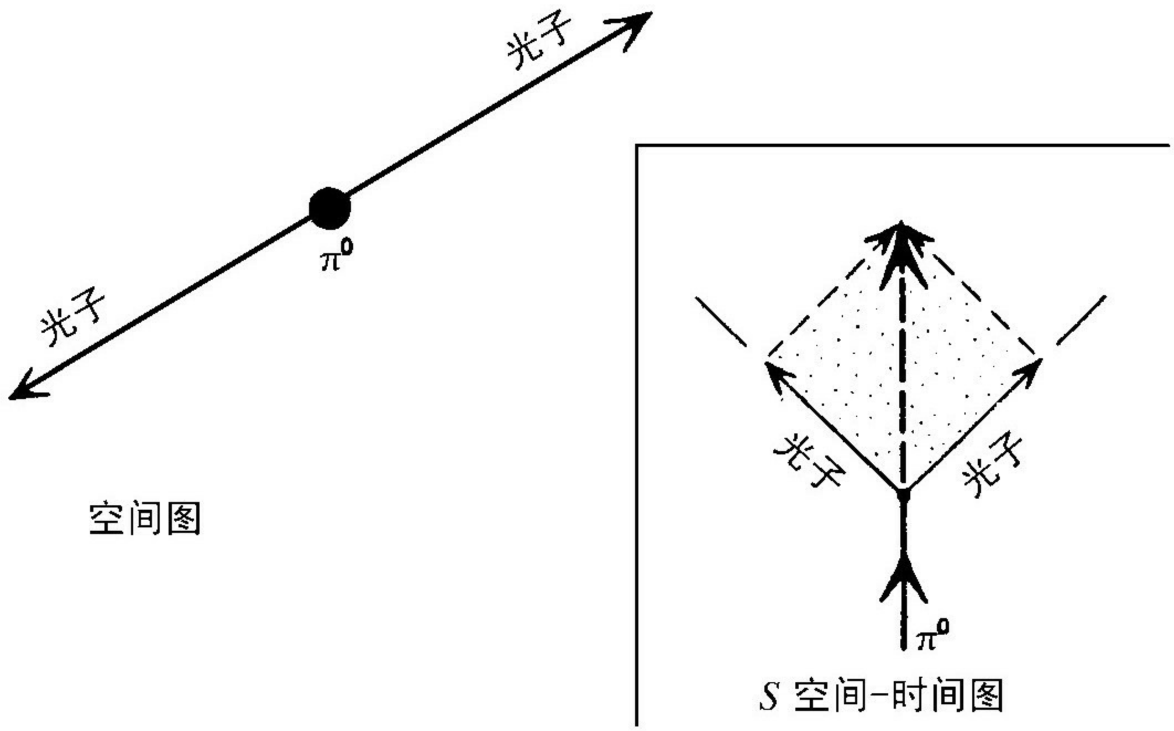


图5.35　一个有质量的π^0 介子衰变为两个零质量的光子。从空间——时间图可以看出能量——动量的四维矢量是守恒的：按照平行四边形加法定律（阴影所示），π^0 介子四维矢量是两光子四维矢量之和

值得花时间将我们得到的观点表述得更明白一些。取代质量作用的守恒量是称为能量——动量四维矢量的整体。在闵可夫斯基空间中可把它画成从原点O出发的一个箭头（矢量），它指向O点未来光锥的内部（或者在光子的极端的情况下，处于光锥之上）；见图5.36。这个和物体世界线指向一致的箭头包含有能量、质量和动量的所有信息。这样，此箭头端点在某观察者坐标系中测量的“t值”（或“高度”）表示观察者看到的物体的质量（或能量除以c^2　），而动量（除以c）由其空间分量所提供。

这个箭头的闵可夫斯基“长度”是称为静质量的重要的量。它描述和此物体同处静止的观察者所看到的质量，人们也许会采取将此当作“物体的量”的好的量度的观点。然而，它没有可加性：如果一个系统分裂成两半，则原先的静止量并不是结果的两个静质量的和。回想一下π^0 介子衰变的情形。π^0 介子具有正的静质量，而分裂成的每个光子的静质量为零。但是，可加性对于整个矢量（四维矢量）的确成立，我们现在必须在画在图5.36中的矢量加法定律的意义上进行“相加”。现在我们“物质

的量”正是用整个箭头来测量！

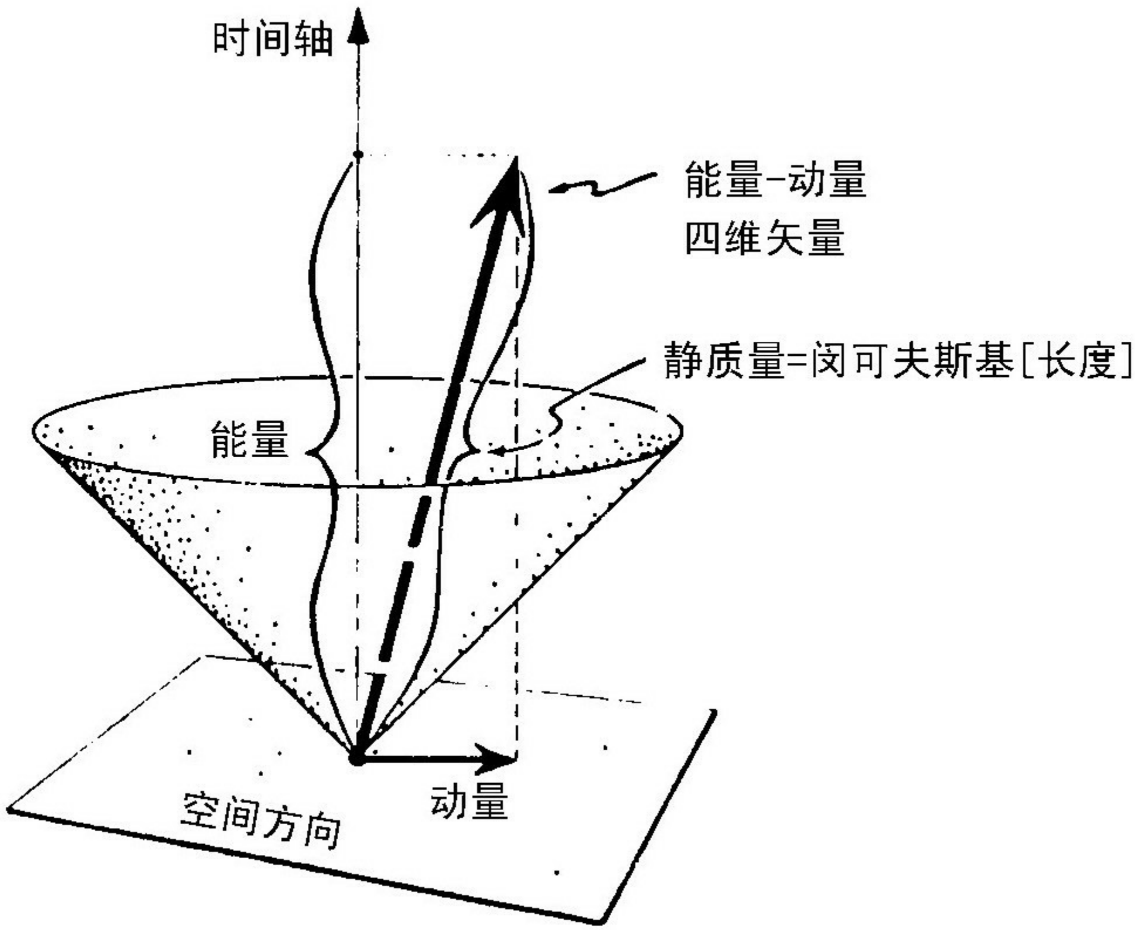


图5.36 能量——动量四维矢量

让我们现在考虑麦克斯韦电磁场。我们知道它携带能量。按照$E=mc^2$，它还应该有质量。这样，麦克斯韦场又是物质！由于麦克斯韦场非常密切地参与到将粒子捆绑在一起的力中，所以这一点我们肯定必须接受。在任何物体中的电磁场一定对其质量有重要的贡献[28]。

关于爱因斯坦引力场又如何呢？在许多方面它和麦克斯韦场很类似。和在麦克斯韦理论中的运动带电体会发射电磁波相似，运动的大质量物体（按照爱因斯坦理论）也会发射出引力波（参阅272页）——它正如电磁波一样以光速传播并携带能量。然而，此能量不是以标准的方式测量的，它是前面讲到的张量能量。在（纯粹）引力波中，此张量实际上处处为零！尽管如此，人们可采用如下观点，空间——时间的曲率

（现在全部由张量外尔给出）多多少少能代表引力波中的“东西”。但是引力能是非定域的，也就是说，人们不能靠考察一个有限区域的空间——时间曲率来决定能量的度量。引力场的能量——并因此质量——的确是非常滑的鳝鱼，我们无法将其钉死在任何清楚的位置上。尽管如此，我们必须严肃认真地对付。它肯定在那里，必须把它考虑在内才能使质量的概念在大范围内守恒。已找到一个可用于引力波的好的（并且是正的）质量测度（*Bondi 1960*，*Sachs 1962* ），但它的非定域性变成这种样子，在两次辐射爆之间的空间——时间的平坦区域中（和飓风眼中的静区很类似）此测度有时为非零，在该处其实完全没有曲率（参阅 *Penrose and Rindler 1986*，*P445* ）（亦即外尔和里奇均为零）！在这种情形下，我们看来不得不做出结论，如果此质量——能量必须存在某处的话，则应该处于这个平坦的空的空间中——一个完全没有任何种类的物质和场的区域。在这些古怪的情形下，我们“物质的量”或者在那里，在此空的区域中的最空虚之处，或者根本那里也不存在！

这看来纯粹是佯谬。然而，它正是我们最好的经典理论——它们也的确是超等的理论——所告诉我们关于世界的“实”物质的性质。按照经典理论，且不必说我们即将探索的量子理论，物质实体比人们所设想的更模糊得多。它的测量——甚至它是否存在——很清楚地依赖于一些微妙的症结，并且不能仅仅定域地确证！如果这种非定域性都使人迷惑不解的话，我们还要准备迎接更大打击的来临。

第6章　量子魔术和量子神秘

哲学家需要量子理论吗

在经典物理中，存在一个“外面的”客观世界，这一点是和常识相符合的。那个世界以一种清晰的、决定性的方式演化着，并受到被精确表达的数学方程的制约。这一点对于麦克斯韦以及爱因斯坦理论，正如对原先的牛顿理论一样都是正确的。物理世界被认为独立于我们而存在；经典世界究竟“是”什么样子不受我们选择去观察它的方法的影响，而且我们的身体和大脑本身也是那个世界的一部分。它们也被认为是按照同等精密和确定的经典方程演化的。不管我们如何觉得我们清醒的意愿在影响着我们的行为，我们的一切行动都被这些方程所决定。

大多数关于实在的性质、我们清醒的知觉以及表观上的自由意志的严肃的 [1] 哲学论证的背景都具有一幅这样的图像。一些人也许会对量子理论——这一事物的基本的，却是使人困扰的理论——也起作用感到不舒服。量子理论是在20世纪最初的25年由于观察到世界的实际行为和经典物理的描述之间的微妙偏差而产生的。对许多人来说，“量子理论”这一术语仅仅是唤起某种“不确定性原理”的模糊概念。该原理禁止我们在粒子、原子或分子的水平上对之进行精确的描述，所能得到的只是随机的行为。实际上，我们将会发现，尽管量子描述和经典物理彻底不同，它却是非常精确的。此外，我们还将看到，尽管一般的观点与它正相反，在粒子、原子和分子的微小的量子水平上不出现随机性——它们决定性地进行演化——概率似乎是通过某种大尺度的、神秘的和我们能意识感觉的经典世界的呈现相关联的作用而产生的。我们必须理解量子理论如何迫使我们改变物理实在的观点。

人们会以为量子和经典理论之间的偏差非常微小，但事实上它们同时又是许多中观物理现象的基础。固态物体之所以存在、物质的强度和物性、化学的性质、物质的颜色、凝固和沸腾现象、遗传的可能性，还有许多其他熟知的性质需要量子力学才能解释。也许还有意识，它是某种不能由纯粹经典理论来解释的现象。我们的精神也许是来源于那些在实际上制约我们居住的世界的物理定律的某种奇怪的美妙特征的性质，

而不仅仅是赋予称之为经典的物理结构的“客体”的某种算法的特征。在某种意义上，这也许就是“为什么”尽管经典宇宙已经是如此的丰富和神秘，作为有情感的生物，我们必须在量子世界，而不是在完全经典的世界中生活。为了诸如我们这样的思维的知觉的生物可由世界物质构成，是否需要一个量子世界？诸如这样的问题似乎更适合于让一心建造一个可供人居住的宇宙的上帝，而不是我们去解答！但是这个问题和我们也有关系。如果意识不可能是经典世界的一部分，那么我们的精神必须以某种方式依赖于对经典物理的特殊的偏离。这就是我在本书中还要考虑的问题。

如果我们要深入钻研一些哲学的主要问题：我们世界如何行为，以及由什么构成“精神”也就是“我们”，则我们的确必须屈服于量子理论，这个最精确也最神秘的物理理论。有朝一日科学将会给我们提供比量子理论更好的，对自然的更深刻的理解。我个人的看法是，甚至量子理论也只是权宜之计，肯定不足以作为我们实际生活其中的世界的完整图画。但这不可作为我们的借口，如果我们想得到某些我们需要的具有哲学洞察力的东西，我们就必须按照已有的量子理论去理解世界图像。

不幸的是，不同的理论家对什么是这个图像的实在持不同（尽管在观察上等效）的观点。以中心人物尼尔斯·玻尔为首的许多物理学家说根本就没有客观的图像。在量子水平上，“外界”没有什么东西。实在多多少少只是在和“测量”结果的关系上才呈现。按照这种观点，量子理论仅仅提供了计算步骤，而不想对世界的实际进行描述。我认为，这种对理论的看法是过于悲观了，而我采用更正面的看法，对量子描述赋予客观的物理实在：量子态。

存在一个非常精确的方程，即薛定谔方程，它为量子态提供了完全决定性的时间演化。但是在随时间演化的量子态和被看到物理世界发生的实际行为之间存在一种非常古怪的东西。只要我们认定“发生”了测量，我们就必须抛弃我们直到该时刻止辛辛苦苦演化来的物理态，而用它来计算该态会“跃迁”到一族新的可能的态的不同的概率。除了这个量子跃迁的怪异之外，对于物理形态还存在什么是裁决“测量”实际上已经进行了的问题。测量装置本身毕竟假定是由量子元件建造的，所以也要按照决定性的薛定谔方程演化。“测量”的实际发生是否必须伴随有意识的存在？我想量子理论家中只有少数人会采取这种观点。大概人类的观察者自身也是由微小的量子元件所组成的吧！

我们将在本章的后面考察量子态“跃迁”的某些奇怪推论——例如，为什么在一处的“测量”似乎会在遥远的区域引起一个跃迁！在这之前，我们还将碰到其他的怪现象：有时一个物体可以分别非常好地通过两个不同的途径。但是一旦同时允许通过两条途径它们就会互相抵消，使得任何一条也通不过！我们还将仔细地考察实际上量子态是如何描述的。我们会看到这种描述和相应的经典描述差别有多大。例如，粒子会一下子在两处出现！当一起考虑几个粒子时，我们会看到量子描述是多么复杂。人们会发现，个别粒子本身并没有单独的描述，而必须考虑所有它们在一道的不同形态的复杂叠加。我们会看到为什么同一类的不同粒子不能有各自的本体。我们将仔细地考察自旋的（基本是量子力学的）古怪性质。我们还将考虑由令人困惑的“薛定谔猫”的理想实验所引发的重要问题以及理论家们提出的试图解决这个基本迷惑的各种不同看法。

本章中的一些材料并不像前面（或后面）章节那么明白易解，有时又有点过于专业性。在描述中我尽量做到诚实，这样我们必须更勤勉一些。其目的在于真正理解量子世界。在论证的不甚清楚之处，我建议你要坚持下去，以期对整个结构有点印象。如果无法完全理解也不必沮丧；它是这个学科本身的性质！

经典理论的问题

我们何以得知经典物理不能真正描述我们的世界呢？主要的理由来自实验。量子理论不是理论家们加在我们身上的预言，大多数理论家是无可奈何地被赶到这一个在哲学的许多方面不满意的、奇怪的世界观上去。其根本的原因在于两种物理现象必须共存：粒子，每一粒子只由很少的有限数目（6）的参数（3个位置和3个动量）来描述；还有场，它需要无限多个参数来描述。这种二分法在物理上不是真正协调的。在粒子和场处于平衡（亦即“完全安置好”）的系统中，所有粒子的能量都会被场抽取走。这即是所谓的“能量均分”现象的结果：系统处于平衡时，能量被公平地分布在所有的自由度上。由于场具有无限多个自由度，所以根本就没有给可怜的粒子留下任何能量！

尤其是，经典原子不能是稳定的，粒子的所有运动都转移到场的波动模式中去。让我们回顾一下伟大的新西兰/英国实验物理学家恩斯特·卢瑟福在1911年引进的原子的太阳系模型。公转的电子处于行星的地

位，中心的太阳为原子核所取代，它们在很微小的尺度上由电磁力而不是引力绑在一起。一个基本的，并且似乎是不可逾越的问题是，当一个公转电子绕着核子时，按照麦克斯韦理论应发射出电磁波，其强度在比1秒钟短得非常多的时间间隔里迅速地增强到无穷，同时它以螺线形的轨道向内撞到核上去！然而，人们从未观测到过这类事。在经典理论的基础上理解所观察到的结果是非常困难的。原子会发射出电磁波（光），但是只能以突发的形式，它具有非常特别的分立频率，这就是被观察到的狭窄的光谱线（图6.1）。而且这些频率满足“莫名其妙”的规则[2]，这从经典理论观点看来毫无根据。

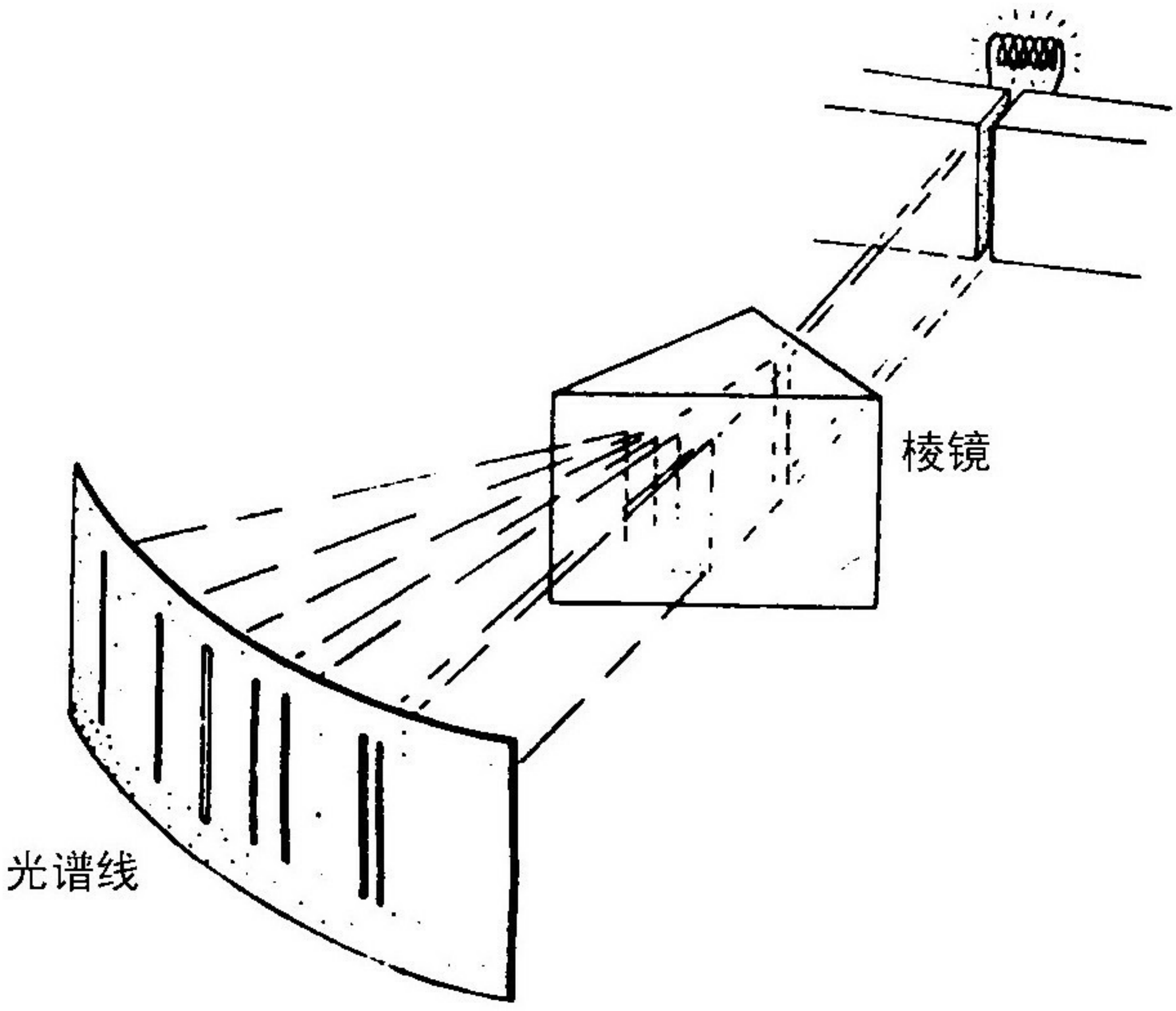


图6.1　经常发现从灼热的物质中的原子发射出的光独具非常特别的频率，可用棱镜把这不同的频率分解，从而提供了原子的特征光谱线

另一种场和粒子不能共存的不稳定性的呈现是称为“黑体辐射”的现

象。想象具有某个确定温度的物体，电磁辐射和粒子处于平衡状态。1900年，瑞利和金斯计算出，所有能量都会被场吸收光——没有极限！此处发生了物理上荒谬的事情（“紫外灾难”：能量不断地跑到场中去，跑到越来越高的没有上限的频率上去），而自然本身却更谨慎。在场振动的低频处，能量正如瑞利和金斯所预言的那样。但是在预言到灾难的高端，实际观察显示，能量分布并没有无限增加，而是随着频率增加而下落。在给定的温度下，能量的最大值发生在非常特别的频率（也即颜色）处（图6.2）。（火钳的红颜色和太阳的黄——白热实际上是两个人们所熟知的例子。）

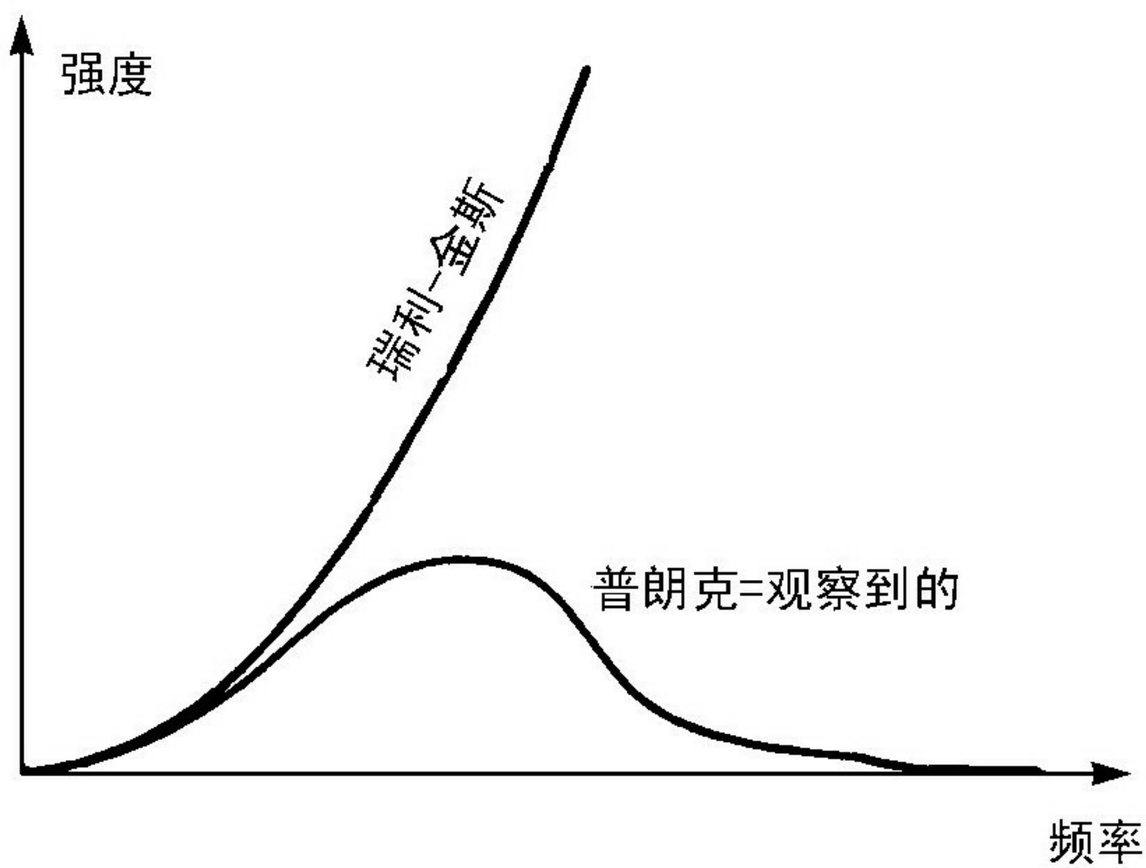


图6.2　经典计算（瑞利——金斯）和观察到的热体（“黑体”）辐射强度之间的偏差导致了普朗克开创的量子理论

量子理论的开端

这些迷惑如何得到解决呢？牛顿原先的粒子理论肯定需要麦克斯韦

场来补充。人们是否可以走到另一极端，假定任何东西都是场，而粒子只是某种场的有限尺度的“结”？这本身也有困难，因为这样的话，粒子可连续地改变它们的形状，可以用无限多不同的方式蠕动和振动。而所有这些我们都没看到。物理世界中的所有种类的粒子都显得是等同的。例如，两个电子完全是相互一样的。甚至原子和分子只能采用分立的不同的形态 [3] 。如果粒子是场的话，那么需要一些新的因素去使场采取分立的特征。

1900年，才华横溢的，但又是保守谨慎的德国物理学家马克斯·普朗克提出了一个革命性的思想用以压制“黑体”的高频率的模式：电磁振动只能以“量子”的形式发生，量子的能量E和频率ν之间有一确定的关系：

$E=h\nu$，

h为一自然的基本常数，现在被称作普朗克常数。令人叹为观止的是，普朗克利用这个荒谬绝伦（无法无天）的因素，能够在理论上得到和观察一致的作为频率函数的强度，这就是现在所谓的普朗克辐射定律。（按日常标准来看，普朗克常数是非常小的，大约为6.6×10^{-34} 焦耳秒。）普朗克凭此壮举揭示了量子理论光临的曙光。尽管在爱因斯坦提出另一个使人惊愕不已的设想，即电磁场只能以这种分立的单位存在之前，普朗克理论并没有引起多大注意。我们记得，麦克斯韦和赫兹指出了光是由电磁场的振荡所组成的。这样一来，按照爱因斯坦——以及牛顿在两个多世纪以前所坚持的——光本身实际上应为粒子！（在19世纪初叶，卓越的英国理论家兼实验家托马斯·杨显明地建立了光为波动的事实。）

光如何由粒子又同时由场振荡所组成的呢？这两个概念的矛盾似乎是不可调和的。某些实验事实很清楚地显示光是粒子，而另一些事实则指出光为波动。1923年，法国贵族及富有洞察力的物理学家路易·德布罗意王子在他的博士论文中（该论文是爱因斯坦认可的！）使这个粒子——波动的图像更加混淆，他提出物体的粒子本身有时应像波动那样行为！任何质量为m的粒子的德布罗意波频率ν也满足普朗克关系式。这与爱因斯坦的$E=mc^2$ 相结合，即告诉我们ν和m之间的关系是：

$h\nu=E=mc^2$ 。

这样，按照德布罗意的设想，自然不遵循作为经典理论特征的粒子和场的二分法！事实上，任何以某频率ν振荡的东西都只能以分立的单位质量$h\nu/c^2$ 发生。自然以某种方式设计建造一个协调的世界，在其中粒子和场振动被认为是同一东西！或者，在她的世界中包含某种更微妙的要素，而“粒子”和“波动”两词汇只不过传达了它部分的合适的图像。

1913年，丹麦物理学家及20世纪主要科学思想家尼尔斯·玻尔再次极其漂亮地利用了普朗克关系。一个绕核公转的电子角动量（参阅215页）只能为h/2π的整数倍，这即是玻尔规则。后来狄拉克为了省事引进了符号$\hbar$：

$$\hbar = h/2\pi。$$

这样，绕着任何轴的角动量的可允许值为

$$0，\hbar，2\hbar，3\hbar，4\hbar，\cdots$$

原子的“太阳系模型”在加上这个新的要素后，就得到了在相当的准确度上，自然实际服从的许多分立的稳定的能量级和谱频率的“怪异的”规则。

玻尔漂亮的设想虽然极其成功，却只是提供了称为“旧量子论”的某种临时的“凑合物”的理论。我们今天所知道的量子理论是由后来的两套独立的方案所产生的。它们是由两个杰出的物理学家所开创的：一位是德国的威尔纳·海森伯，另一位是奥地利的厄文·薛定谔。这两种方案（分别为1925年的“矩阵力学”和1926年的“波动力学”）在初始时显得完全不同，但是很快发现它们是等同的，并且很快就被包摄到一个更合理更一般的框架中去。这个框架是在不久之后首先由英国伟大的理论物理学家保罗·阿得林·毛里斯·狄拉克提出。我们将在以下几节了解该理论的概要以及它的非同寻常的含义。

双缝实验

让我们考虑这一“原型的”量子力学实验。一束电子或光或其他种类的“粒子——波”通过双窄缝射到后面的屏幕去（图6.3）。为了确定起见，我们用光做实验。按照通常的命名法，光量子称为“光子”。光作为粒子（亦即光子）的呈现最清楚地发生在屏幕上。光以分立的定域性的能量单位到达那里，这能量按照普朗克公式E=hν恒定地和频率相关。从未接收过“半个”或任何部分光子的能量。光接收是以光子单位的完全有或完全没有的现象。只有整数个光子才被观察到。

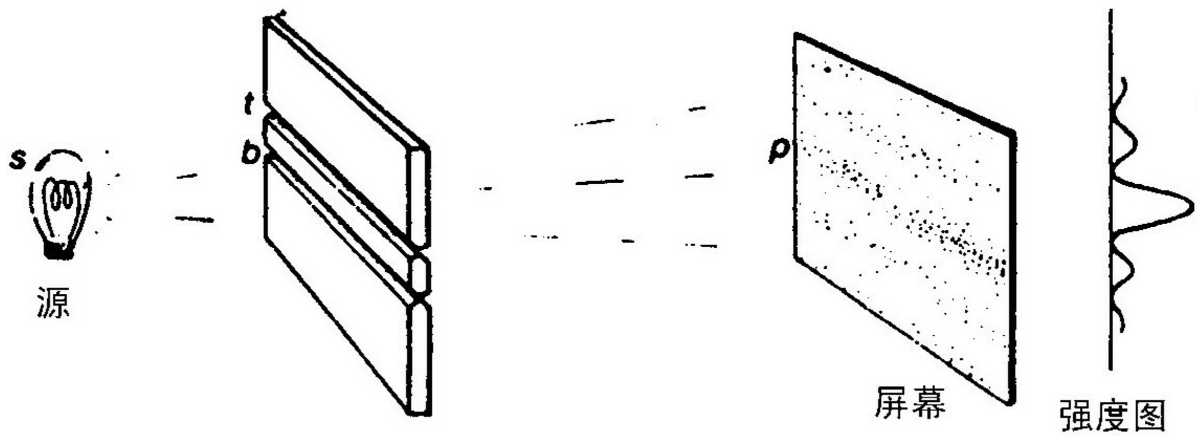


图6.3　单色光的双缝实验

然而，光子通过缝隙时似乎产生了类波动的行为。先假定只有一条缝是开的（另一条缝被堵住）。光通过该缝后就被散开来，这是被称作光衍射的波动传播的一个特征。但是，这些对于粒子的图像仍是成立的。可以想象缝隙的边缘附近的某种影响使光子随机地偏折到两边去。当相当强的光也就是大量的光子通过缝隙时，屏幕上的照度显得非常均匀。但是如果降低光强度，则人们可断定，其亮度分布的确是由单独的斑点组成——和粒子图像相一致——是单独的光子打到屏幕上。亮度光滑的表观是由于大量的光子参与的统计效应（图6.4）。（作为比较：一个60瓦的电灯泡每1秒钟大约发射出100000000000000000000个光子！）光子在通过狭缝时的确被随机地弯折——弯折角不同则概率不同，就这样得到了所观察到的亮度分布。

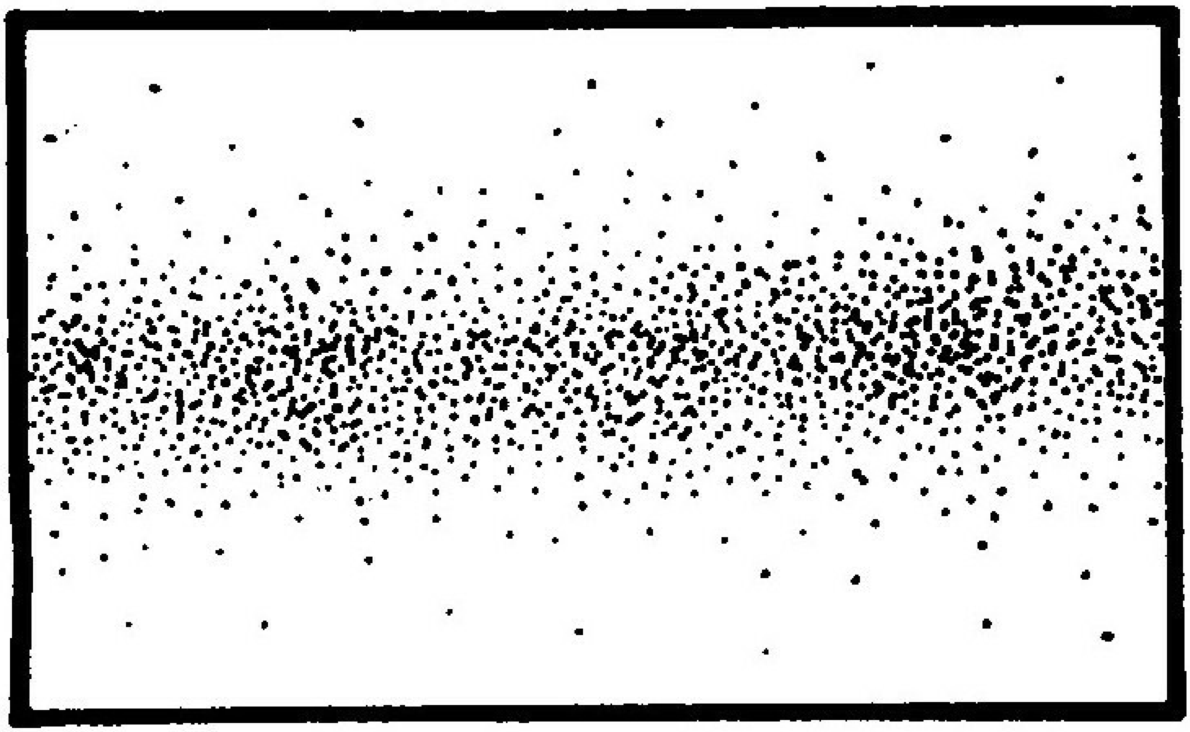


图6.4　只有一个缝隙打开时屏幕上的强度模式——分立小点的分布

然而，当我们打开另一条缝隙时就出现了粒子图像的关键问题！假设光是来自于一个黄色的钠灯，这样它基本上具有纯粹的非混合的颜色——用技术上的术语称为单色的，也即具有确定的波长或频率。在粒子图像中，这表明所有光子具有同样的能量。此处波长约为5×10^{-7} 米。假定缝隙的宽度约为0.001毫米，而且两缝相距0.15毫米左右，屏幕大概在1米那么远。在相当强的光源照射下，我们仍然得到了规则的亮度模式。但是现在我们在屏幕中心附近可看到大约3毫米宽的称为干涉模式的条纹的波动形状（图6.5）。我们也许会期望第二个缝隙的打开会简单地把屏幕的光强加倍。如果我们考虑总的照度，这是对的。但是现在强度的模式的细节和单缝时完全不同。屏幕上的一些点——也就是模式在该处最亮处——照度为以前的四倍，而不仅仅是二倍。在另外的一些点——也就是模式在该处最暗处——光强为零。强度为零的点给粒子图像带来了最大的困惑。这些点是只有一条缝打开时粒子非常乐意来的地方。现在我们打开了另一条缝，忽然发现不知怎么搞的光子被防止跑到那里去。我们让光子通过另一条途径时，怎么会在实际上变成它在任何一条途径都通不过呢？

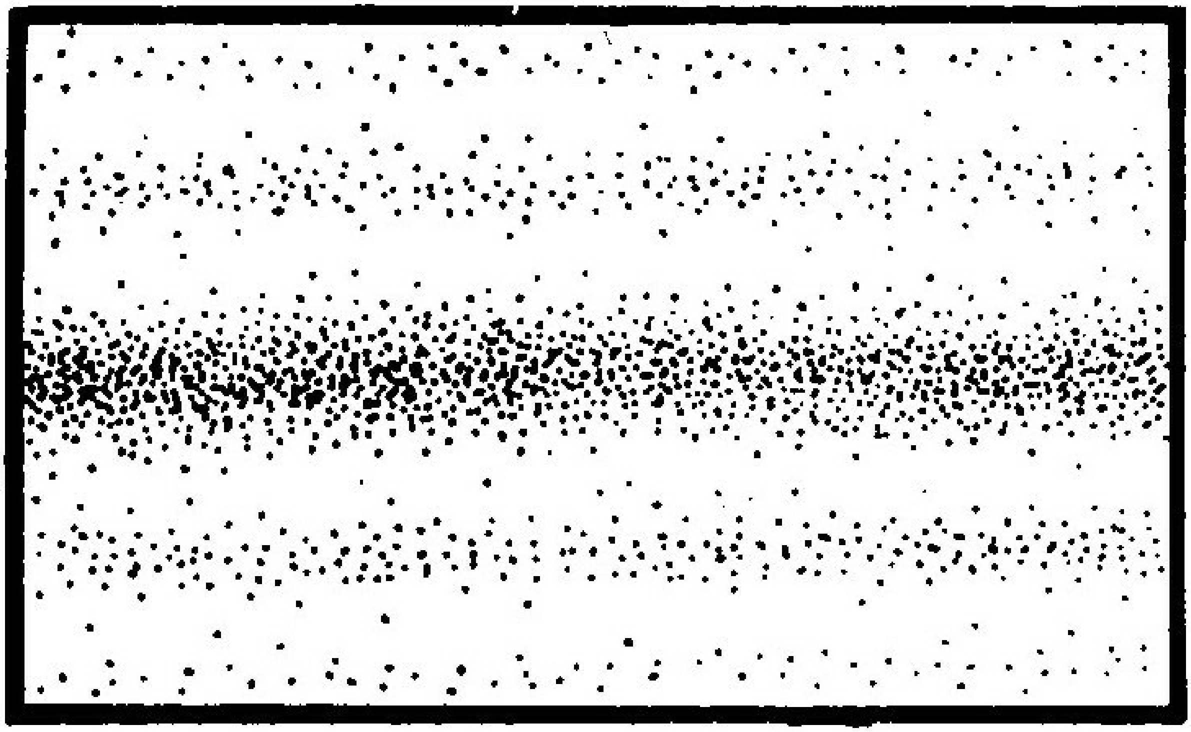


图6.5　两个缝隙同时打开时屏幕上的强度模式——分立小点的波动状分布

在光子的情形下，如果我们取它的波长作为其“尺度”的度量，则第二条缝离开第一条缝大约有300倍“光子尺度”那么远（每一条缝大约有两个波长宽）（图6.6），这样当光子通过一条缝时，它怎么会知道另一条缝是否被打开呢？事实上，对于“相消”或者“相长”现象的发生，两条缝之间的距离在原则上没有受到什么限制。

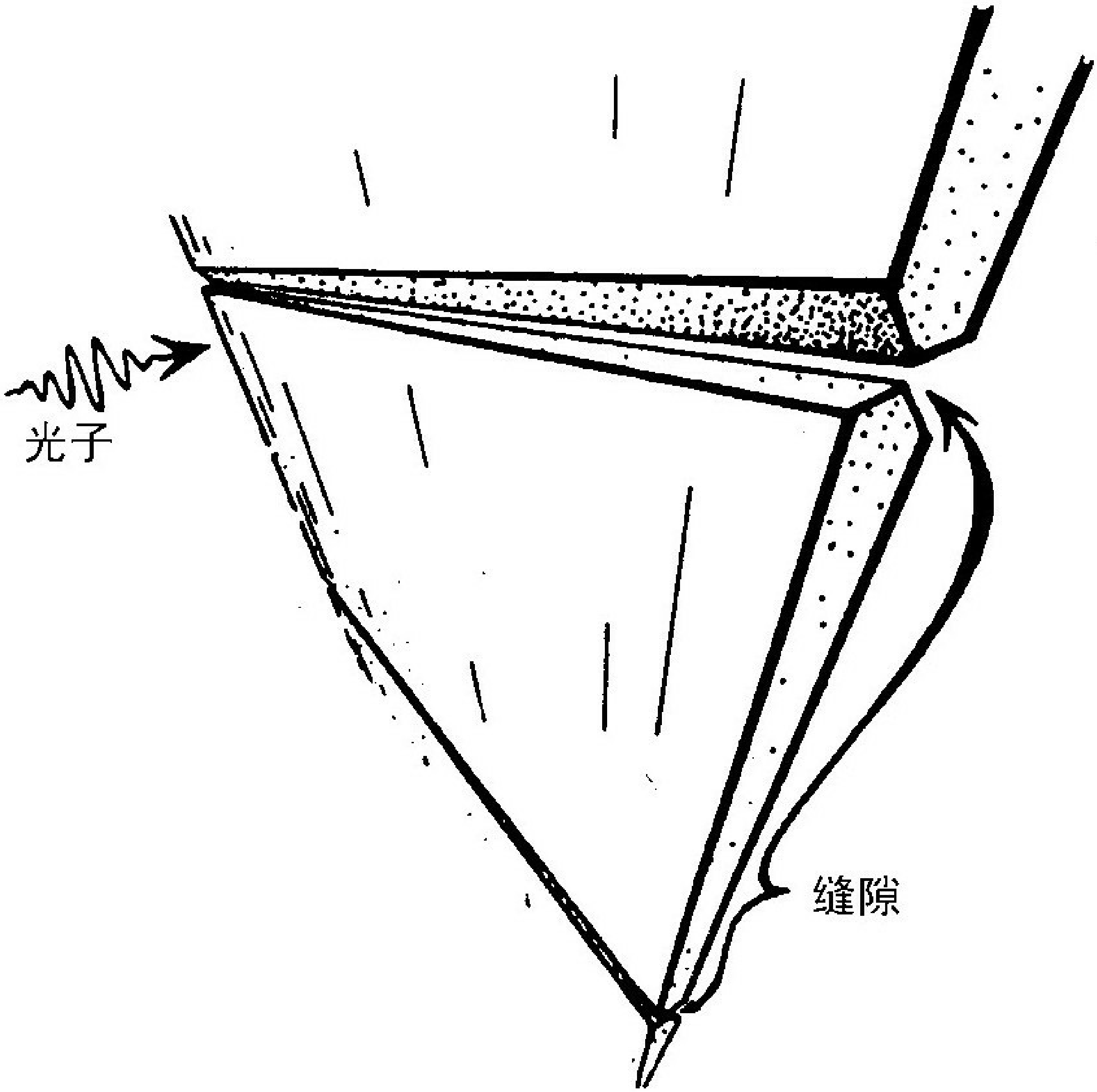


图6.6　从光子的观点看缝隙！大约在300倍“光子尺度”外的第二条缝是开还是闭，对它而言怎么会有影响呢？

当光通过缝隙时，它似乎像波动而不像粒子那样行为！这种抵消——相消干涉——是波动的一个众所周知的性质。如果两条路径的每一条分别都可让光通过，而现在两条同时都开放，则它们完全可能会相互抵消。我在图6.7中解释了何以至此。如果从一条缝隙来的一部分光和从另一条缝隙来的“同相”（也就是两个部分波的波峰同时发生，波谷也同时发生），则它们将互相加强。但是如果它们刚好“反相”（也就是一个部分波的波峰重叠到另一部分的波谷上），则它们将互相抵消。在双

缝实验中，只要屏幕上到两缝隙的距离之差为波长的整数倍的地方，则波峰和波峰分别在一起发生，因而是亮的。如果距离差刚好在这些值的中间，则波峰就重叠到波谷上去，该处就是暗的。

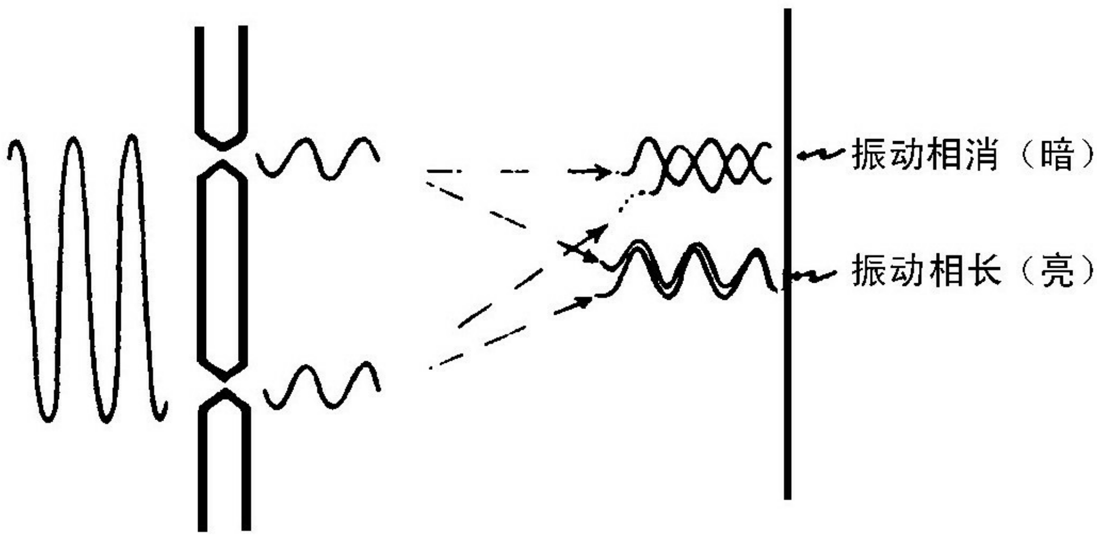


图6.7 在纯粹波动图像中，我们可按照波动的干涉来理解屏幕上亮的和暗（虽然不是分立）的模式

关于通常宏观的经典波动同时以这种方式通过两个缝隙没有任何困惑之处。波动毕竟只是某种媒质（场）或者某种包含有无数很小点状粒子的物体的一种“扰动”。扰动可以一部分通过一条缝隙，另一部分通过另一条缝隙。但是这里的情况非常不同：每一个单独光子自身是完整的波动！在某种意义上讲，每个粒子一下通过两条缝隙并且和自身干涉！人们可将光强降得足够低使得保证任一时刻不会有多于一个光子通过缝隙的附近。相消干涉现象，因之使得两个不同途径的光子互相抵消其实现的可能性，是加在单独光子之上的某种东西。如果两个途径之中只有一个开放，则光子就通过那个途径。但是如果两者都开放，则两种可能性奇迹般地互相抵消，而发现光子不能通过任一条缝隙！

读者应该深入思考一下这一个非同寻常事实的重要性。光的确不具有有时像粒子有时像波那样的行为。每一个单独粒子自身完全地以类波动方式行为；一个粒子可得到的不同选择的可能性有时会完全相互抵消！

光子是否在实际上分成了两半并各自穿过一条缝隙呢？大多数物理学对这样的描述事物的方式持否定态度。他们坚持说，两条途径为粒子

开放时，它们都对最后的效应有贡献。它们只是二中择一的途径，不应该认为粒子为了通过缝隙而被分成两半。我们可以考虑修正一下实验，把一个粒子探测器放在其中的一条缝隙，用来支持粒子不能分成两部分再分别通过两缝隙的观点。由于用它观测时，光子或任何其他种类的粒子总是作为单独整体而不是整体的一部分而出现，我们的探测器不是探测到整个光子，就是根本什么也没探测到。然而，当把探测器放在其中的一条缝隙处，使得观察者能说出光子是从哪一条缝隙通过时，屏幕上的波浪状的干涉花样就消失了。为了使干涉发生，显然必须对粒子“实际上”通过那一条缝隙“缺乏知识”。

为了得到干涉，两个不同选择都必须有贡献，有时“相加”——正如人们预料的那样相互加强到两倍——有时“相减”——这样两者会神秘地相互“抵消”掉。事实上，按照量子力学的规则，所发生的事比这些还更神秘！两种选择的确可以相加（屏幕上最亮的点），两者也的确可以相减（暗点）；但它们实际上也会以另外奇怪的组合形式结合在一起，例如：

“选择A”加上i乘以“选择B”，

这儿i是我们第3章的“负一的平方根”（$=\sqrt{-1}$）（在屏幕上中等强度的地方）。事实上任何复数都能在“不同选择的组合”中起作用！

读者可能会记得在第3章时我的复数对于“量子力学的结构是绝对基本的”警告。这些数绝不仅仅是数学的精巧。它们通过令人信服的、使人意外的实验事实来迫使物理学家注意。我们必须接受复数权重才能理解量子力学。现在我们接着考虑它的推论。

概率幅

在上面的描述中利用光子并无任何特别之处。这里可以同样好地利用电子或任何其他种类的粒子或者甚至原子。量子力学的规则坚持，甚至连棒球和大象都应以这种古怪的方式行为，不同选择的可能性可用复数的组合“相加起来”！然而，我们从未在实际中看到棒球或大象这种奇怪方式的叠加。为什么我们没有见到呢？这是一个困难的富有争议的问题，我现在还不想去对付之。作为工作规则，现在让我们简单地假设物理描述有两种不同可能的水平，我们将其称为量子水平和经典水平。我

们只在量子水平上利用这些古怪的复数组合。棒球和大象是经典水平上的对象。

量子水平就是分子、原子和亚原子粒子的水平。这通常被认为是非常“小尺度”现象的水平，但是这个“小”实际上并非是指物理尺度。我们将会看到量子效应能在许多米甚至1光年的距离上发生。如果认为只牵涉到非常小的能量差，这才有点接近于认为某种东西是“处于量子水平上”的特征。（以后我将尽力弄得更精确些，尤其是在第8章的465页。）经典水平就是我们直接了解的“宏观”水平。在这水平上，我们的“事物”发生的通常图像是正确的，并且可以使用通常的概率观念。我们将看到在量子水平上，我们必须使用的复数和经典概率有紧密的关系。它们并不真正相同，但是为对付这些复数，先回顾一下经典概率的行为是有益的。

考虑一个不确定的经典情形，两种选择之中我们不知哪一种会发生。可将这种情形描述作这些选择的“加权”组合：

p×“选择A”加上q×“选择B”，

此处p为A发生的概率，而q是B发生的概率。（要记住，概率是在0和1之间的实数。概率1表明“一定发生”，而概率0表明“一定不发生”。概率1/2表明“发生和不发生是同等可能的”。）如果A和B是仅有的不同选择，则两者概率的和必须是1：

p+q=1.

然而如果还有其他选择，则此和可以比1小。那么，比率p：q就给出了发生A和发生B的概率的比率。在只有两种选择时，发生A和发生B的实际概率分别为p/（p+q）和q/（p+q）。如果p+q比1大，我们还可以这样解释。（这可能是有用的，例如，只要我们进行了多次的实验，p为发生A的次数，q为发生B的次数。）如果p+q=1，我们就说p和q是归一化的，这样它们就给出了实际的概率，而不仅仅是概率的比率。

在量子力学中我们将做一些显得与此非常相似的事，现在p和q变成为复数——我将使用w和z分别表示之：

w×“选择A”加以z×“选择B”。

我们如何解释w和z呢？由于它们会各自独立地变为负数或者复数，它们肯定不是通常的概率（或概率比），但是在许多方面很像概率。它们被叫作（适当地归一化之后——见后面）概率幅，或简单地称作幅度。此外，人们经常用这类暗示概率的术语，如：“发生A的幅度为w和发生B的幅度为z。”它们不是实际的概率，但是我们假装它们是——或宁愿说成概率在量子水平上的相似物。

通常的概率如何起作用呢？考虑一个宏观对象将有助于理解，譬如说打一个球使之穿过两个洞中的一个再到后面的屏幕去——正如上述的双缝实验那样（图6.3），但现在我们用经典的宏观球取代了前面讨论的光子。从s将球打到上洞的概率为P（s, t），打到下洞的概率为P（s, b）。而且，如果我们在屏幕上选取特定的一点p，只要球的确通过t，则到此特定的p点的概率为P（t, p），而球通过b到达p的概率为P（b, p）。如果只有上面的洞t是开放的，则球通过t到达p的实际概率为将从s到t的概率乘上从t到p的概率：

P（s, t）×P（t, p）.

类似地，如果只有下面的洞是开放的，则球从s到p的概率为：

P（s, b）×P（b, p）.

如果两个洞都开放的话，则从s通过t到达p的概率仍为第一表达式P（s, t）×P（t, p），正如只有t洞开放时那样。而从s通过b到p的概率仍为P（s, b）×P（b, p）。所以，从s到p的总概率P（s, p）为两者之和：

P（s, p）=P（s, t）×P（t, p）+P（s, b）×P（b, p）.

在量子水平上，除了现在是奇怪的复的幅度起着我们前面的概率的作用外，其规则和这一模一样。这样，在上面考虑的双缝实验中，光子从源s到上缝t我们有一幅度A（s, t），从上缝到达屏幕上p点有一幅度A（t, p），两者相乘得到从s通过t到达p的幅度：

A（s, t）×A（t, p）.

作为概率，假定上缝是开的，不管下缝是否打开，这都是正确的幅度。类似地，假定b是开的，则存在光子从s通过b到达p的幅度（不管t是

否打开）：

A（s, b）×A（b, p）.

如果两条缝隙都打开，我们可得到光子从s到p的总幅度：

A（s, p）=A（s, t）×A（t, p）+A（s, b）×A（b, p）.

这一切都非常好。但是，我们在量子效应被放大达到经典水平从而知道如何去解释这些幅度之前，它对我们并没有多大用处。我们可把一个光子探测器或光电管放在p处，它提供了把量子水平的事件——光子抵达p—放大成经典的可辨别得出的发生，例如听得见的“咔嗒”一声。（如果屏幕的作用相当于照相底版，使得光子留下可见的斑点，那么这也是一样的。但为了清楚起见我们就用光电管好了。）必须存在产生“咔嗒”一响的实际的概率，而不仅仅是这些神秘的“幅度”！当我们从量子水平变到经典水平时，如何从幅度过渡到概率呢？人们发现这里有一种非常美丽而神秘的规则。

其规则是我们必须对量子的复的幅度取平方模以得到经典的概率。什么是“平方模”？回忆一下我们对复平面上的复数的描述（第3章118页）。复数z的模|z|简单地就是z离开原点（也就是点0）的距离。平方模 $|z|^2$ 即是这个数的平方。这样，如果我们写：

z=x+iy，

这儿x和y都是实数。由于从0到z的连线为直角三角形O， x，z的斜边，从勾股定理得知我们所需的平方模是：

$|z|^2 = x^2 + y^2$.

注意，为了使之成为一个真正的“归一化的”概率，$|z|^2$ 的值必须在0和1之间。这表明对于适当归一化的幅度，在复平面上z必须处于单位圆内的某处（图6.8）。然而，有时我们要考虑组合：

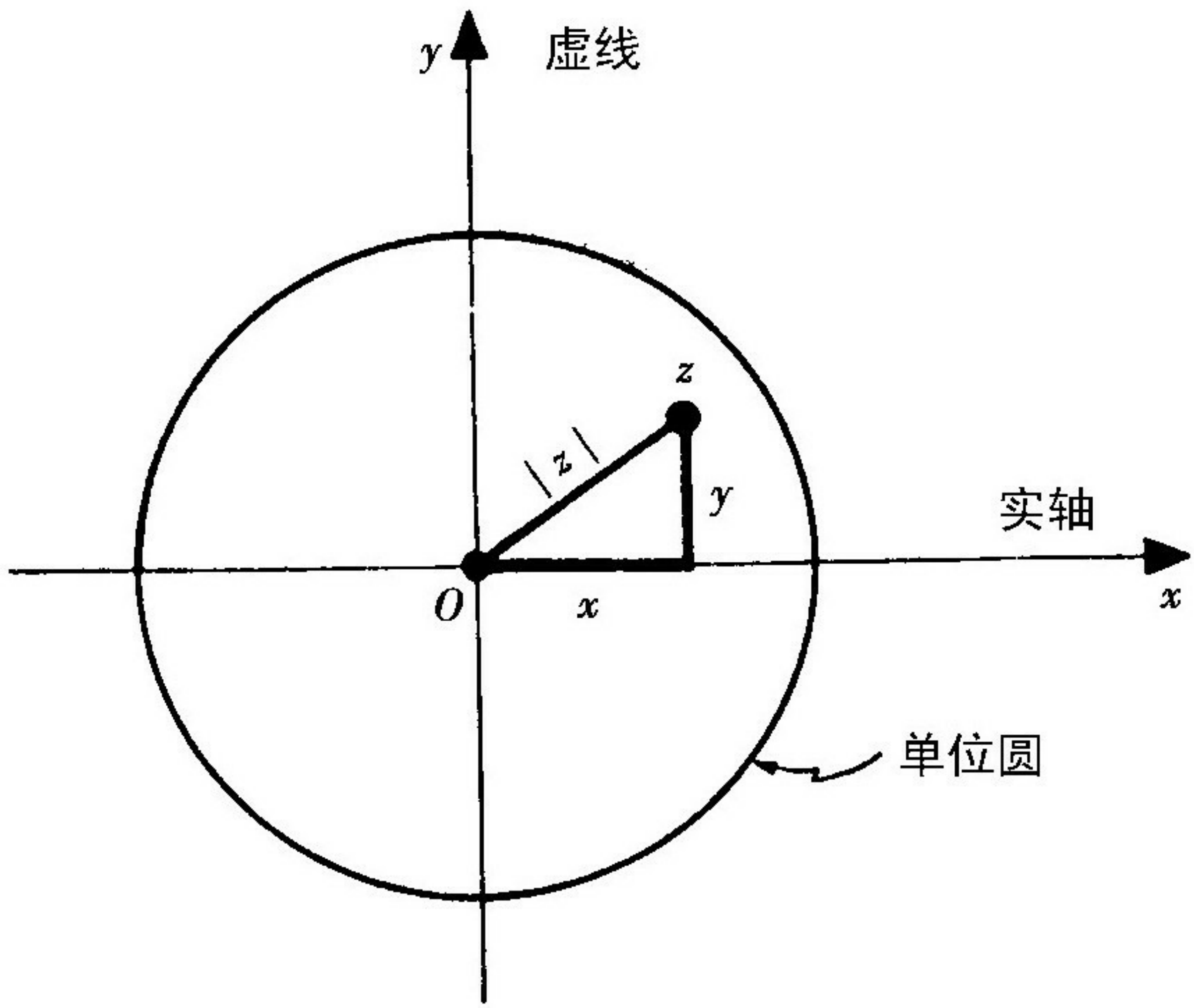


图6.8　用复平面上单位圆内的点z来代表概率幅。其与中心的距离的平方$|z|^2$可成为当效应被放大到经典水平时的实际概率

w×“选择A”+z×“选择B”，

此处w和z仅仅是和概率幅成比例，它们没必要在单位圆内部。它们归一化（并因此提供真正的概率幅）的条件是平方模的和必须为1：

$$|w|^2+|z|^2=1。$$

如果它们不是归一化的，则A和B的实际幅度应分别为：

$$w/\sqrt{|w|^2+|z|^2} \text{ 和 } z/\sqrt{|w|^2+|z|^2},$$

它们都处于单位圆内部。

现在我们看到，概率幅根本不像真正的概率，而更像概率的“复数平方根”。当量子水平上的效应被放大到经典水平上时，这会产生什么影响呢？我们记得，在进行概率和幅度运算时，我们有时要将它们相乘，有时将它们相加。第一点值得注意的是，乘法运算在从量子过渡到经典规则时没有什么问题。这是因为乘积的模数等于各自模数的乘积的这一显著的数学事实：

$$|zw|^2=|z|^2|w|^2。$$

（这个性质可由第3章的一对复数的乘积的几何描述立即得出；但是若按照实部和虚部z=x+iy　w=u+iv，这还算是一点奇迹。不妨试一下！）

此事实的含义是，如果只有一条通道对粒子开放，也就是在双缝实验中只有一条缝隙（譬如t）开放，即可以“经典地”论证，不管是否在中间某点（譬如在t）进行附加的粒子检测，出来的概率必须是一样的[44]。我们可以在两个阶段或只在最后取平方模，也即：

$$|A(s,t)|^2\times|A(t,p)|^2=|A(s,t)\times A(t,p)|^2,$$

对于最后的概率，其结果都是一样的。

然而，如果多于一条通道可让粒子通过（也即如果两条缝隙都开放的话），则我们要求和，而量子力学的特征就在这里开始出现。当我们取两个复数w和z的和（w+z）的平方模时，通常不能得到它们各自的平方模的和；还有附加的“修正项”：

$$|w+z|^2=|w|^2+|z|^2+2|w||z|\cos\theta 。$$

此处θ为点z和点w对复平面原点所张的角（图6.9）。（我们知道，一个角的余弦是一直角三角形的“邻边/斜边”比。不熟悉上式的敏捷读者可用第3章引进的几何去直接推导之。实际上，这正是众所周知的“余弦规则”，只不过稍微伪装了一下！）正是修正项2|w||z|cosθ提供了量子力学的不同选择间的量子干涉。cosθ的值的范围在-1和1之间。我们在θ=0°时有cosθ=1。这时这两种选择相互加强，使得总概率比单独概率之和更大。我们在θ=180°时有cosθ=-1，这时这两种选择便相互抵消，使得总概率比单独概率之和更小（相消干涉）。我们在θ=90°时有cosθ=0。这时得到了一种中间状态，两种概率相加。对于大的或复杂的系统修正项通常被“平均掉了”——因为cosθ的“平均”值为零——我们就余下通常的经典概率的规则！但是在量子水平上这些项提供重要的干涉效应。

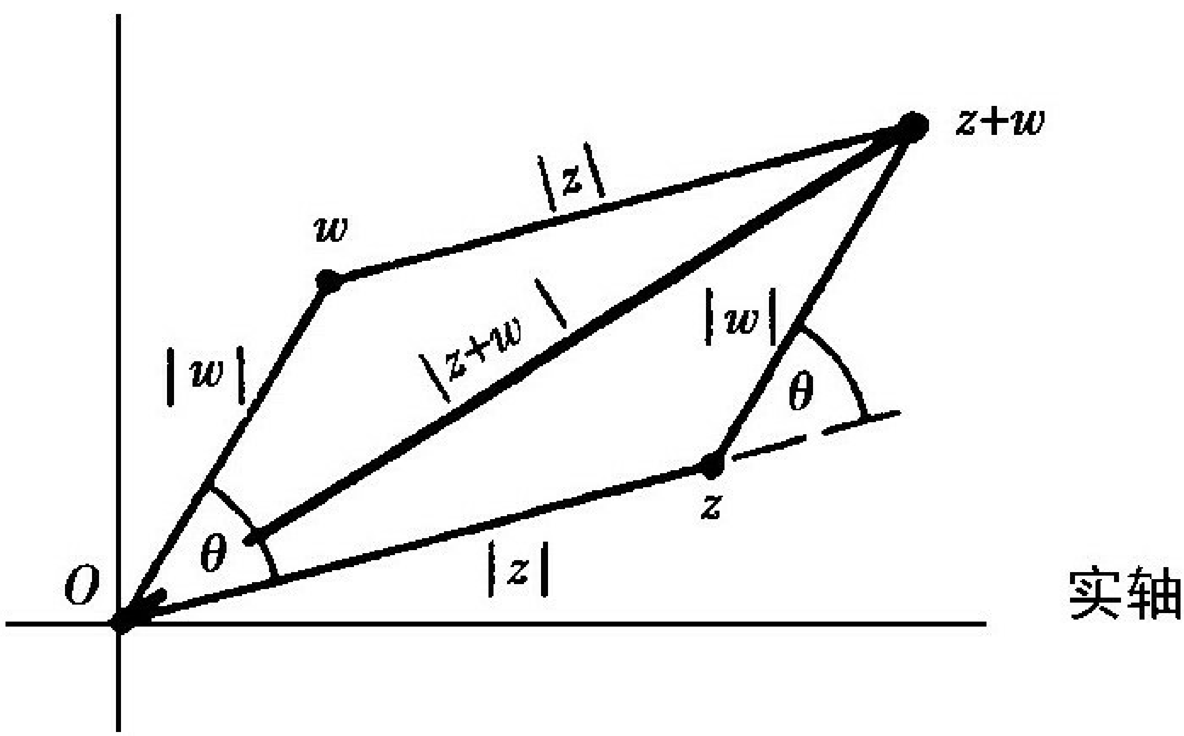


图6.9　有关两个幅度的和的平方模的修正项2|w||z|cosθ的几何

考虑双缝都打开时的双缝实验。到达p的光子幅度为和w+z，此处

w=A（s, t）×A（t, p）和z=A（s, b）×A（b, p）。

在屏幕的最亮的点我们有w=z（这样cosθ=1），所以

$$|w+z|^2=|2w|^2=4|w|^2$$

为只有一条缝开放时概率$|w|^2$ 的4倍——所以当光子数很大时光强变大到4倍，这与观察相一致。在屏幕的暗的点我们有w=-z（这样cosθ=-1），所以

$$|w+z|^2=|w-w|^2=0,$$

也就是零（相消干涉！），又与观察相一致。在刚好中间的点我们有w=iz或w=-iz（这样cosθ=0），所以

$$|w+z|^2=|w\pm iw|^2=|w|^2+|w|^2=2|w|^2$$

给出只有一条缝的强度的两倍（这是经典粒子的情形）。我们在下一节的结尾处会看到如何去实际计算亮、暗和中间的位置。

还有最后一点必须加以评论。当双缝都开放时，通过t到达p的粒子的幅度确是w=A（s, t）×A（t, p），但是我们不能将其平方模$|w|^2$ 当作粒子“实际”通过上面的缝隙而到达p的概率。这会导致没有意义的答案，特别是如果p是在屏幕上的暗的地方时。但是，如果我们决定“检测”光子是否在t存在，把它在那儿的存在（或不存在）的效应放大到经典的水平，则可用$|A(s, t)|^2$ 作为光子实际到达t的概率。但是这样的检测抹去了波浪状的模式。为了使干涉发生，我们必须保证光子在通过缝隙时仍维持在量子水平上，以使得两个不同途径能共同有贡献并且有时会互相对消。单独的选择途径只有幅度，而没有概率。

粒子的量子态

这些在量子水平上为我们提供了“物理实在”的什么图像呢？在这里，一个系统的不同的“选择可能性”必须一直共存，并且用奇怪的复数权重加在一起。许多物理学家本身对是否能找到这样的图像感到绝望。相反的，他们断言，他们喜欢量子力学仅仅是它为我们提供了计算概率的步骤，而不是物理世界的客观图像的观点。有些人断定量子理论不可能有客观图像——至少没有一种和物理事实相一致。我认为这样的悲观主义是没有根据的。在我们已经讨论到的基础上，采取这种看法无论如何都是不成熟的。我们将在下面讨论某些量子效应更令人吃惊的困惑，进而更全面地了解这种绝望的原因。但是，现在我们暂且更乐观地前进，并接受量子力学告知我们所必须面临的情景。

这就是一种量子态所呈现的图像。我们现在考虑一个单独的量子粒子。一个粒子由它的空间位置经典地决定。为了知道它下一步还要做什么，我们还需要知道它的速度（或等效地，它的动量）。在量子力学中，粒子所能到达的每个单独位置都是它所能得到的一个“选择”。我们看到所有的选择必须以复数的权重组合在一道。这一复权重的集合描述了粒子的量子态。标准的做法是用希腊字母Ψ（发“psi”的音）表示权重的集合，Ψ被认为称作粒子的波函数的位置的复函数。对于每一位置x，波函数都有一个用Ψ（x）表示的特殊的值，它是粒子处于x的幅度。我们可用单独的Ψ来表示整个量子态。我所采取的观点是，粒子所处位置的物理实在的确是它的量子态Ψ。

我们如何画出复函数Ψ呢？一下子将所有的三维空间都画出是有点困难，所以我们先简化一些并假定粒子被限制在一维的线上——譬如说沿着标准（笛卡儿）坐标系的x轴上。如果Ψ是一个实函数，则我们可以想象和x轴垂直的“y轴”并画出Ψ的图[图6.10（a）]。但是，为了描述复函数的Ψ的值，我们在这儿需要一个“复的y轴”——它必须是一个复平面。我们在想象中可以利用空间的两个维：譬如把空间的y方向当作复平面的实轴，z方向作为虚轴。我们可以把Ψ（x）画成在这个复平面[也即是通过x轴上每一点的（y，　z）平面]上的一点，这样就可得到一个波函数的精确的图像。这一点随着x的变化而变化，而它的轨迹在空间画出一条绕着x轴附近的曲线[见图6.10（b）]。我们称这条曲线为粒子的Ψ曲线。如果在一指定点x处放置一台粒子检测器，则在该点找到该粒子的概率可由幅度Ψ（x）取平方模而得到：

$|\Psi(x)|^2$。

这正是Ψ曲线离开x轴的距离的平方[45]。

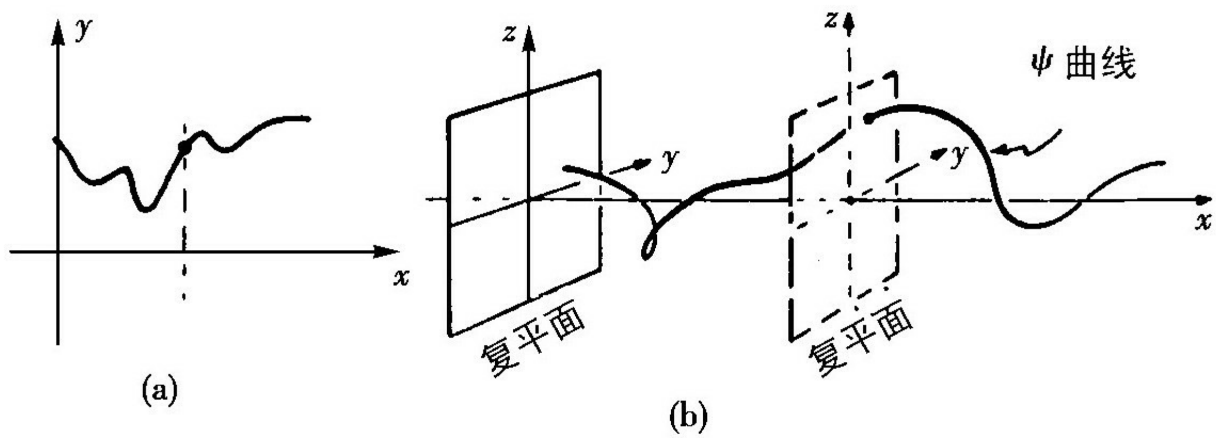


图6.10（a）实变量x的实函数的图（b）实变量x的复函数Ψ的图

为了画出在所有三维物理空间上波函数的完整的图，五维是必须的：三维是物理空间，加上画出Ψ（x）的复平面的二维。然而，我们简化了的图仍是有助的。如果我们选择沿着物理空间的任一特别的线来考察波函数，我们就可简单地让x轴沿着这线，并临时利用其他两个空间方向来提供所需的复平面。这对理解双缝实验是有用的。

正如我前面提到的，在经典物理中为确定粒子下一步怎么走，人们需要知道它的速度（或动量）。在这里，量子力学以显著的经济的方式为我们提供了这些。波函数Ψ中已经包含有不同可能动量的各种幅度！（一些不满的读者考虑到我们已经将点粒子的简单的经典图像变复杂了这么多，也许认为现在该是有一点经济的“时候”了！虽然我非常同情这种读者，我得警告他们赶紧将扔给他们的这一些先捡起来，因为后面还有更坏的来临！）如何从Ψ来决定速度幅度呢？实际上考虑动量幅度更好。（我们记得动量是速度乘以粒子的质量，215页。）人们所做的是把所谓的调和分析应用到函数Ψ上去。我不可能在这里仔细地解释它。但它和处理乐声有紧密的关系。任何波形都能被分解成为不同“和声”的和（这就是“调和分析”术语之来源）。它们是不同音调（亦即不同频率）的纯净的乐音。在波函数Ψ的情形，“纯粹乐音”对应于粒子可能有不同的动量，而每一“纯粹乐音”对Ψ贡献的大小提供了该动量值的幅度。而“纯粹乐音”本身被称作动量态。

动量态在Ψ曲线上看起来是什么样子的呢？它看起来像个螺旋，其正式的数学名字叫螺旋线（图6.11）[46]。卷得紧的螺旋对应于大动量，而几乎不卷的只具有很小的动量。极限情形是根本不卷，而Ψ曲线变成直线：这是零动量的情形。这里隐含有著名的普朗克关系。卷得紧表明短波长和高频率，并因此高动量和高能量；而卷得松表明低频率和低能量，能量E总是和频率ν成比例（E=hν）。如果复平面以正常的方法指向，亦即上面给出的按照右手定则的x，y，z描述，那么在x轴正方向上的动量对应于右旋的螺旋（这正是通常用的螺旋）。

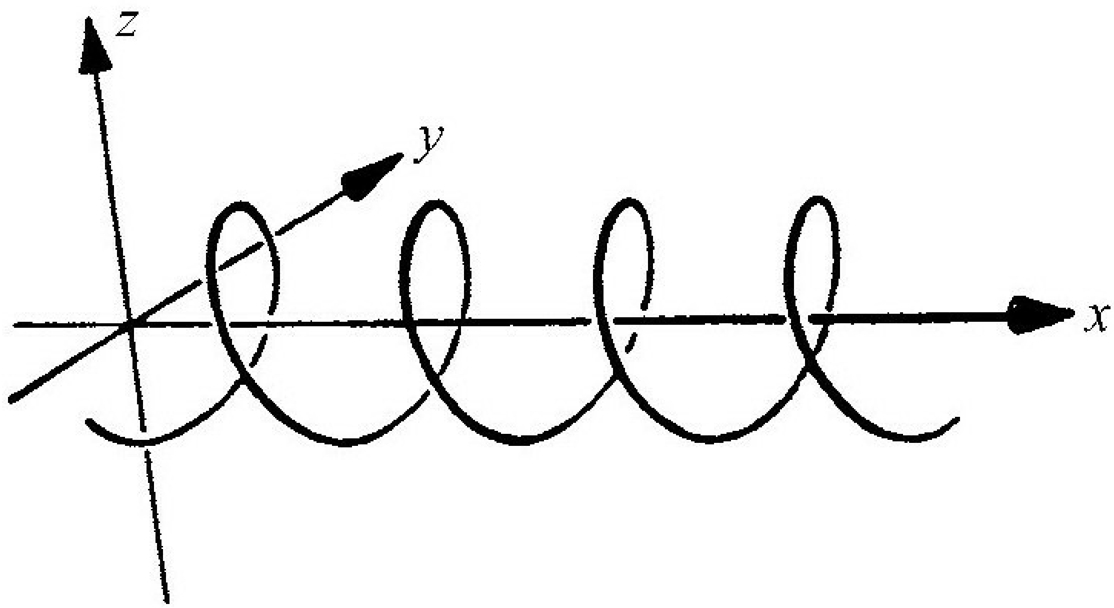


图6.11　动量态具有螺旋形状的Ψ曲线

不像上面那样按照通常的波函数，而是按照动量的波函数来描述量子态有时更有用。这归结为把Ψ按照不同的动量态而展开，从而建立一个新的函数$\tilde{\varphi}$。这回它是动量p而不是位置x的函数。它的值$\tilde{\varphi}$ p）对于每一个p给出了p动量态对Ψ的贡献的大小。（p空间称作动量空间。）$\tilde{\varphi}$的解释是，对于每一特别选定的p，复数$\tilde{\varphi}$（p）给出粒子具有动量p的幅度。

在函数Ψ和$\tilde{\varphi}$之间的关系有一个数学术语。这些函数称为相互的傅里叶变换——这是以法国工程师兼数学家约瑟夫·傅里叶（1768—1830）命名的。在此我只对该关系做些评论。第一点是在Ψ和$\tilde{\varphi}$之间存在一个显著的对称。我们可以应用在本质上和从Ψ得到$\tilde{\varphi}$的同样的步骤

从Ψ得到$\tilde{\varphi}$ 。现在是对$\tilde{\varphi}$ 进行调和分析。而“纯粹乐音”（也就是在动量空间表像中的螺旋）被称作位置态。每一位置x在动量空间决定一这样的“纯粹乐音”，而这个“纯粹乐音”对$\tilde{\varphi}$ 的贡献的大小决定了Ψ（x）的值。

一个位置态本身在通常的位置空间表像中对应于在一个给定的x值处的非常尖锐的峰，除这一点外任何位置的幅度都为零。这种函数称作（狄拉克）δ函数——尽管由于它在x处的值为无限，从而它在技术上并不是通常意义上的“函数”。同样地，动量态（也即位置表像空间中的螺旋）在动量空间表像中给出δ函数（图6.12）。这样，我们看到了螺旋的傅里叶变换是一个δ函数，而且反之亦然！

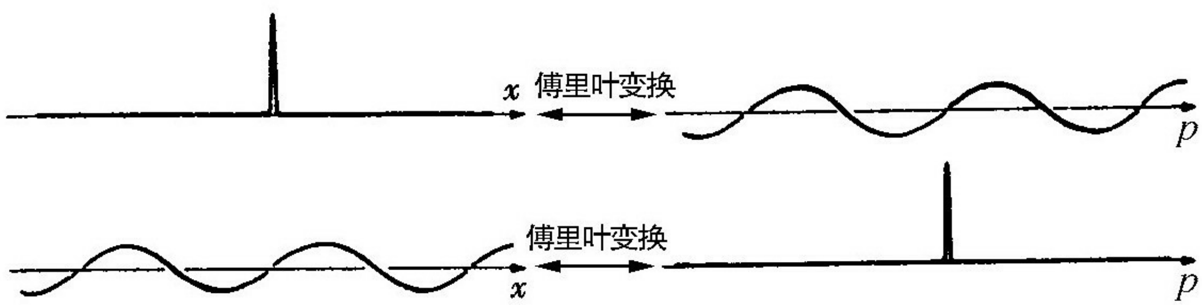


图6.12　位置空间中的δ函数变换成动量空间中的螺旋，反之亦然

每当人们测量粒子的位置，位置空间的描述总会用到。这种测量归结于做一些事情，将不同可能的粒子位置的效应放大到经典的水平。（粗略地讲，光电管和照相底版进行了光子位置的测量。）每当人们测量粒子的动量，动量空间的描述总会用到，这种测量就是将不同的可能的动量的效应放大到经典的水平（反冲效应或晶体的衍射可用于动量测量）。在每种情形下，相应的波函数（Ψ和）的平方模给出了所要测量结果的所要的概率。

在本节结束之前我们再一次回到双缝实验。我们已经知道，按照量子力学，甚至一个单独的粒子都应像波动一样行为。这个波动为波函数Ψ所描述。动量态是最“类似波动”的波。我们在双缝实验中摹想具有确定频率的光子；这样光子的波函数是由在不同方向的动量态组成。这些态中的螺旋的螺矩都是相同的，这螺矩又称作波长。（波长由频率所固定。）

每个光子波函数一开始从源S散开来并且通过两个缝隙（在缝隙上不做任何检测）而到屏幕上去。只有波函数的一小部分从这缝隙出来。我们将每一条缝隙当作从该处分别散开来的波函数的新源。这两部分波函数互相干涉。这样，当它们到达屏幕时，在有些地方互相叠加，在另外一些地方互相抵消。为了找到它们在何处叠加和何处抵消，我们在屏幕上取点p并考察其到两条缝隙t和b的直线，沿着tp有一个螺旋，沿着bp有另一个螺旋。（我们沿着st和sb也有螺旋，但是假定光源到每一条缝隙的距离相同，则在缝隙处两个螺旋刚好旋转了一样多。）现在，当这些螺旋到达屏幕的p点处旋转了多少得由直线tp和bp的长度决定。当这些长度的差为波长的整数倍时，则两个螺旋在p点就从它们的轴向同一方向位移（亦即θ=0°，这儿的θ的意思和上节一样），这样相应的幅度就互相叠加，我们得到一个亮点。当这些长度的差为波长的整数倍加上半波长时，则两个螺旋在p点从它们的轴向相反方向位移（θ=180°），这样相应的幅度就互相抵消，我们得到一个暗点。在所有其他情形下，这两个螺旋到达p时位移间有某一角度，这样幅度就以某种中间的方式相加，我们得到中等的光强（图6.13）。

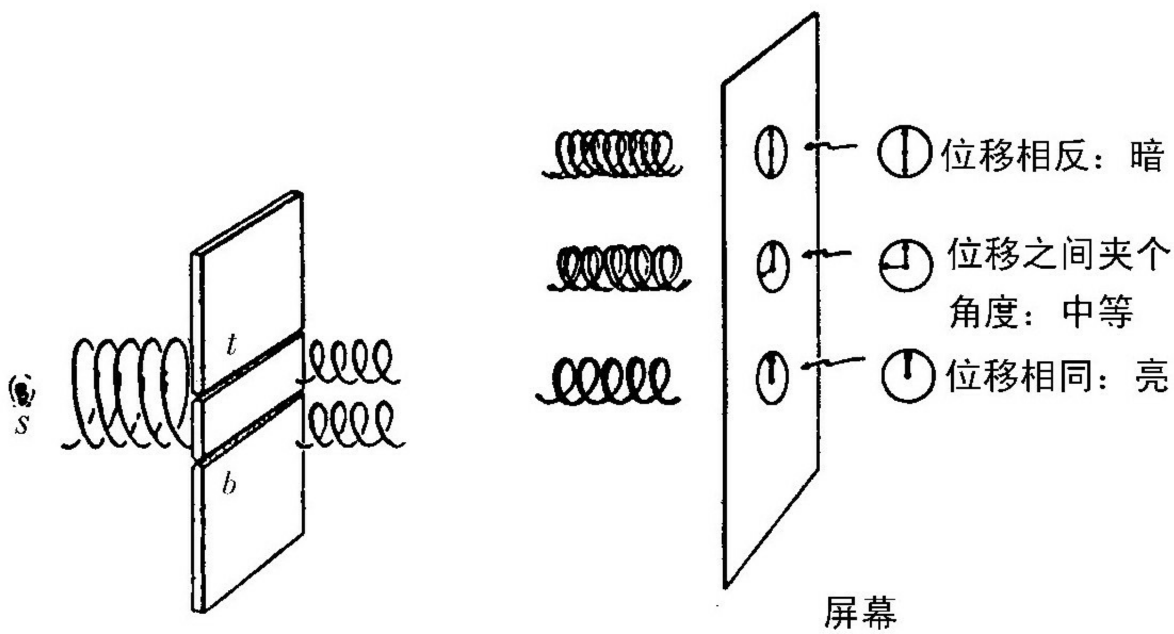


图6.13　按照光子动量态螺旋的描述来分析双缝实验

不确定性原理

大多数读者都听说过海森伯的不确定性原理。根据这一原理，不可能同时将一个粒子的位置和动量精确地测量（亦即放大到经典的水平）。更糟糕的是，这些精度，譬如分别为Δx和Δp的乘积有一绝对极限，它由下面的关系式给出：

$$\Delta x \Delta p \geqslant \hbar \text{。}$$

这一公式告诉我们，位置x测量得越准确，则动量p的测量就越不准确，反之亦然。如果位置被测量到无限精确，则动量就变得完全不确定；另一方面，如果动量被精确地测量，则粒子的位置就变得完全不确定。为了从海森伯关系给出的极限大小得到一些感性认识，假定将一个电子的位置测量到纳米（10^{-9} 米）的精度，那动量会变得这样的不确定，以至于人们不能预料1秒钟之后电子是否离其不到100千米！

一些描述使人相信，似乎这仅仅是测量过程中固有的粗陋。相应地根据这种观点，在刚才考虑的电子的情形下，为了找到它的位置不可避免地赋予了它这等强度的“随机的反冲”，使得电子以海森伯原理所表明的数量级的巨大的速度冲撞。人们在其他的描述中认为不确定性是粒子自身的一个性质，它的运动有一种固有的随机性，这表明在量子水平上它的行为是内在的不可预见的。还有另一种说法认为，量子粒子是某种不可理喻的东西，对此经典位置和动量的概念均不适用。我对这几种看法都不喜欢。第一种有点误导，第二种肯定是错的，而第三种过于悲观。

波函数的描述究竟告诉了我们什么？首先让我们回忆一下动量态的描述。这是动量被准确指定的情况。Ψ曲线为一个螺旋，它离开轴的距离一直是一样的。所以不同位置的幅度都具有相同的平方模。如果要进行位置测量的话，则在任何一点找到该粒子的概率和在任何其他地方一样。粒子的位置是完全不确定的！关于位置态又如何呢？现在Ψ曲线是一个δ函数，位置被精确地固定在δ函数的尖峰处——其他地方的幅度均为零。在动量空间表像中最容易得到动量幅度。现在Ψ曲线为一个螺旋，而不同动量的幅度具有相等的平方模。在测量粒子动量时，其结果会变得完全不确定！

考察位置和动量都只被部分地限制的中间情形是有趣的，只要它们和海森伯关系相一致就可以了。图6.14画出了这种情形的Ψ曲线和相应的$\tilde{\Psi}$ 曲线（相互的傅里叶变换）。我们注意到只在非常小的范围

内每一曲线到轴的距离明显地不为零。曲线在远处非常紧密地环抱着轴。这样，不管是在位置空间还是在动量空间中都只有在一个非常有限的区域平方模才有可觉察到的大小。因此，粒子在空间可以相当定域，但有一定的弥散，类似地，动量也是相当确定，粒子以相当确定的速度运动，而可能的粒子位置的弥散不随时间增加太大。这样的粒子态被称作波包，经常将它作为一个经典粒子的量子论的最好近似。但是动量（或速度）值的弥散表明波包将随时间弥散。原先开始的位置越定域，则弥散开得越快。

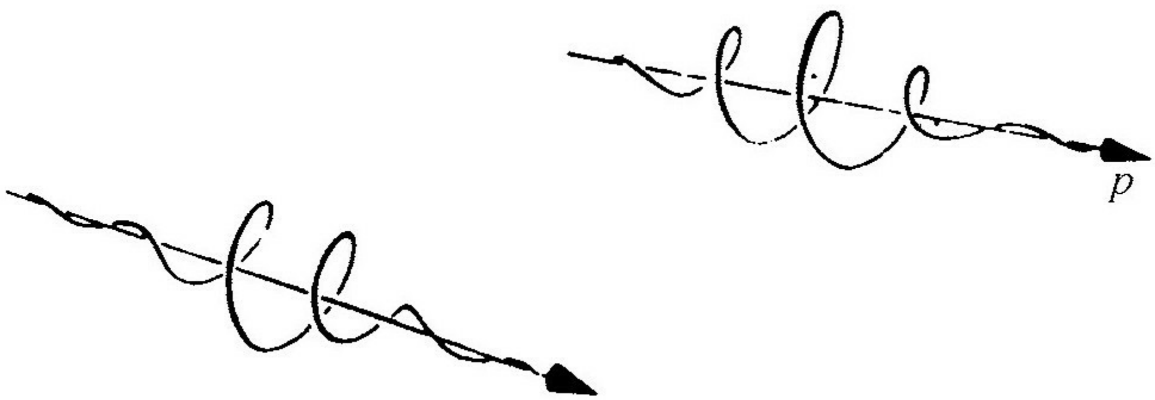


图6.14　波包。这些波包在位置空间和动量空间中都是定域的

U和R演化步骤

在描述波包的时间发展中隐含着薛定谔方程，它告诉我们波函数在时间中的实际演化。薛定谔方程实际上是说，如果我们将Ψ分解成动量态（“纯粹乐音”），那么每一个单独的分量将以问题中具有此动量的经典粒子速度去除c^2 而得到的速度离开。薛定谔数学方程在实质上是以更加紧凑的形式写下这些。下面我们再看它的精确形式。它有点像哈密顿或麦克斯韦方程（和两者有紧密关系）。和那些方程一样，一旦波函数在某一时刻定好，则给出它的完全确定的演化！（参阅369页。）

我们如果将Ψ当作“世界实在”的描述，只要Ψ是由决定性的薛定谔演化所制约，就根本不存在被认为是量子力学固有的特征的不决定性。

让我们将这种演化过程称为U。然而，只要我们“进行一次测量”，将量子效应放大到经典水平，我们就改变了规则。现在我们不用U，而是用完全不同的我称作R的步骤，取量子幅度的平方模以得到经典概率[4]！正是步骤R也只有R在量子理论中引进了不确定性和概率。

决定性的过程U似乎是做量子理论工作的物理学家关心的主要部分；而哲学家则对非决定性的态矢量减缩R（或者，正如有时形象化描述的：波函数的坍缩）更感兴趣。我们是否简单地将R认为是关于一个系统的“知识”的改变，还是认为（正如我认为）是“真正地”发生了什么。我们的确得到了物理系统的态矢量随时间变化的两种完全不同的数学方式。U是完全决定性的，而R是概率定律；U保持量子复叠加原则，但是R显著地违反之；U的作用是连续的，而R公然是不连续的。按照量子力学的标准过程，不存在以任何方式将R“归结”为U的复杂的情况的含义。它干脆是和U不同的过程，提供了量子力学的另一“半”的解释。所有的非决定性都是从R而不是从U来的。为了使量子理论和已有的观测事实美妙地协调，U和R两者都是需要的。

让我们回到波函数Ψ上来。假定它为一个动量态。只要此粒子不和任何东西相互作用，它就会在其余的时间里快乐地维持在那个动量态上。（这是薛定谔方程告知我们的。）无论我们什么时候去“测量其动量”都会得到同一确定的答案。此处不存在概率。和经典理论一样，可预言性在这里是非常清楚的。然而，假定在某一个阶段我们胆敢去测量（也就是放大到经典水平）粒子位置，这回我们就得到了一系列的概率幅，我们必须将它们平方求模。那时候有许许多多的概率，完全无法肯定测量会产生什么结果，其不确定性和海森伯原理相一致。

另一方面，让我们假定Ψ从一个位置态开始（或几乎为一个位置态）。现在，薛定谔方程告诉我们，Ψ不再停留在位置态上，它会很快地弥散开来。尽管如此，其弥散的方式完全由此方程所固定。它的行为没有任何不确定性或随机性。原则上存在去检查此事实的实验（下面还要讲到）。但是，如果我们不明智地决定去测量动量，就会发现所有可能的不同的动量值的幅度平方模相等。实验的结果则是完全的不确定性，这又和海森伯原则相一致，而概率是由幅度的平方模给定。

这无疑是非常奇怪和神秘的。但是它不是不可理喻的世界图像。关于这个由许多非常清楚和准确的定律制约的图像还有许多可说的。然

而，关于何时应该祈求随机性的规则R去取代宿命论的U尚没有清楚的规则。“进行一次测量”是什么含义？为何（何时）对幅度平方取模使之“成为概率”？“经典水平”能被量子力学地理解吗？这些都是在本章后面要讨论的深刻的令人困惑的问题。

粒子同时在两处

我在上面的描述中采取了也许比通常的量子物理学家们更“现实”的关于波函数的观点。我采取了单独粒子的“客观实在”的状态的确是由它的波函数所描述的观点。似乎许多人发现这个观点很难以严肃的方式予以坚持。之所以这样的一个原因是，它牵涉到我们认为单独粒子在空间中弥散开来，而不总是集中在单独的点上的事实。对于一个动量态，由于Ψ在整个空间范围内平均地分布，这弥散达到了极端。人们不认为粒子本身发散到空间中去，而宁愿认为位置是完全不确定的。这样，人们关于位置所能说的是粒子在任何一处正和在另一处同样的可能。然而，我们已经看到，波函数不仅提供了不同位置的概率分布；它还提供了不同位置的幅度分布。如果我们知道这个幅度分布（亦即波函数Ψ），则我们从薛定谔方程就知道粒子的态从一个时刻向另一时刻演化的精确方式。为了这样地决定粒子的“运动”（也就是Ψ随时间的演化），我们需要粒子的这一“发散开去”的观点；而如果我们的确采用这个观点，我们就会看到粒子的运动的确是被精确地决定的。如果我们对粒子施加位置测量，那么关于Ψ（x）的“概率观点”就很合适，因为那时仅仅使用Ψ（x）的平方模的形式：$|\Psi(x)|^2$。

看来必须接受这样的粒子图像，它会在空间的大范围内发散开去，并会一直发散到下一次进行位置测量为止。甚至当一个粒子被定域为位置态后，下一时刻就会开始发散开去。动量态似乎难于被接受为一个粒子存在的“实在”图像，但它也许更难被接受为刚穿过双缝出来的双峰态的“实在”图像（图6.15）。在垂直的方向上，波函数Ψ的形式在每一条缝隙处都有尖锐的峰值。该波函数为上缝有峰值的波函数Ψ_t和在下缝有峰值的波函数Ψ_b的和[47]：

$$\Psi(x) = \Psi_t(x) + \Psi_b(x)。$$

如果认为Ψ代表粒子态的“实在”，那么我们必须接受粒子的确同时在两处的图像！基于这一观点，粒子确实同时穿过两条缝隙。

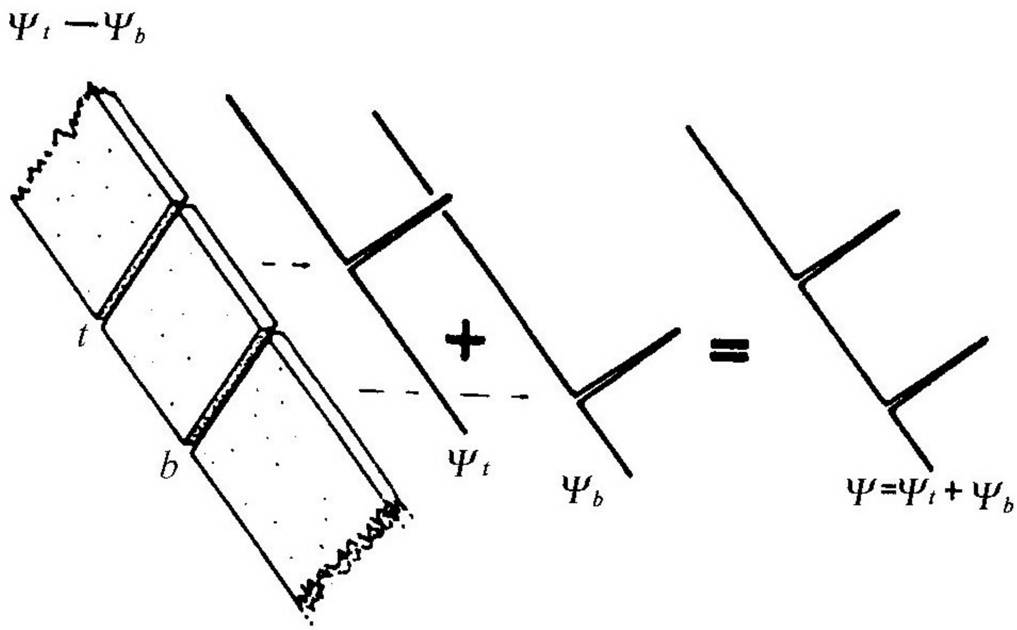


图6.15　当光子波函数从一双缝隙出来时，它同时在两处取得峰值

回忆一下反对粒子“同时穿过两条缝隙”这观点的标准说法：如果我们在缝隙处做测量以确定它是否通过那一条缝隙，我们总是发现整个粒子通过这条或那条缝隙。但是这是因为我们对粒子进行位置测量引起的，这时Ψ仅仅提供和按照平方模步骤一致的粒子位置的概率分布$|\Psi|^2$，而我们的确发现它在这一处或那一处。但是在缝隙处我们还能进行不同于位置测量的其他测量。为此，我们应该知道不同位置x的双缝波函数Ψ，而不仅是$|\Psi|^2$。这样的测量可以将上面给出的双峰态

$\Psi=\Psi_t+\Psi_b$

和另一双峰态，如

$\Psi_t-\Psi_b$

或

$\Psi_t + i\Psi_b$

区别开来。（见图6.16中三种不同情形下的Ψ曲线。）因为确实存在将这些不同可能性区别开来的测量，所以它们必须是光子能存在的不同可能的“实际”方式！

缝隙没有必要靠得很近使“光子”同时穿过它们。为了演示不管它们距离多么远量子粒子总能“同时在两处”，考虑一个稍微和双缝实验不同的实验装置。和以前一样，我们有一个发出单色光的灯泡。每一时刻只发一个光子；但是这回不让光子通过两个缝隙，我们让它从一面倾斜角45°的半镀银的镜面反射出来。（半镀银镜子是一种刚好将射到它上面的光反射一半，而让所余下的一半光直接穿透过去的镜子。）在它遭遇到镜子以后，光子的波函数分裂成两个部分，一部分反射回来，另一部分继续原先光子的方向。波函数又是双峰值的，但是这回双峰是更宽广地分离开了。一个峰描述反射的光子，而另一峰描述透射的光子（图6.17）。此外，两峰的分离随着时间流逝变得越来越大，并随着时间无限地增加。想象波函数的这两部分跑到空间去，而我们整整等待了一年。那么光子波函数的这两部分相距将超过一光年。光子不知怎么搞的发现自己同时出现在相距比一光年还远的两地方！

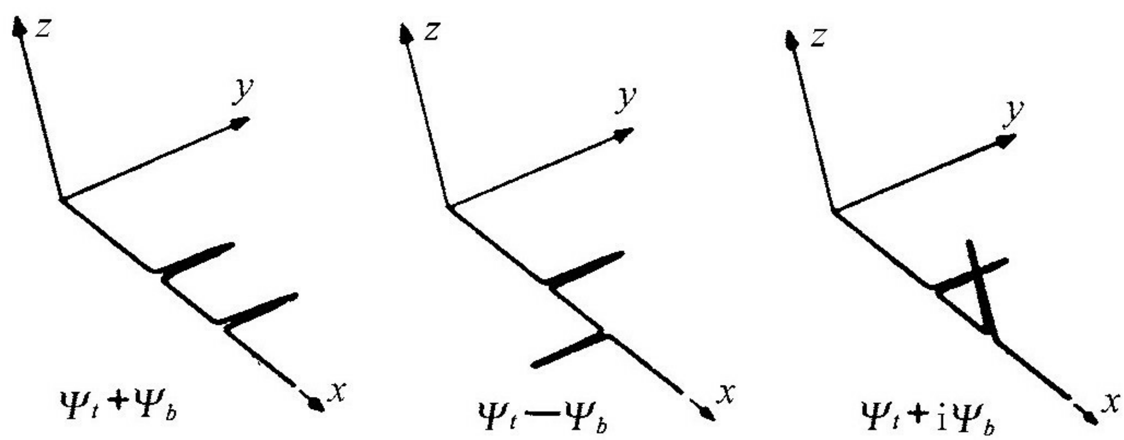


图6.16　三种具有双峰的光子波函数的不同方式

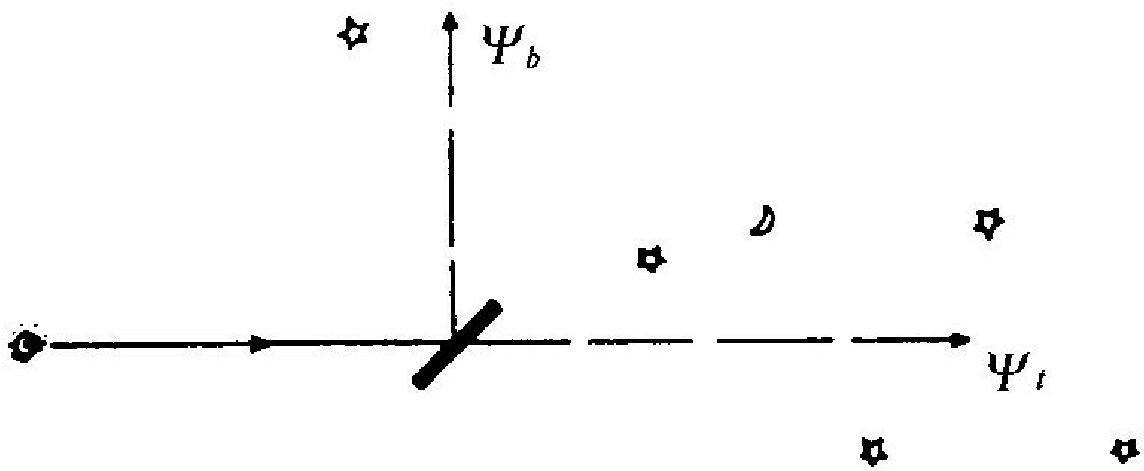


图6.17　双峰波函数的双峰可以分开到一光年那么远。这可以用半镀银镜面做到

是否有理由去认真地接受这样的图像呢？难道我们不能简单地认为光子有百分之五十的机会在一个地方，而另外百分之五十的机会在另一处呢？不，我们不能！不管旅行了多长时间，总能将光束折射回来，使之再互相遭遇，得到两种不同选择的概率权重所得不到的干涉效应。假定光束的两部分各遇到一面全镀银的镜子。我们调整好镜子的角度使之再次遭遇在一起。在交会点放上另一面半镀银镜子，角度刚好和第一面一样。在两束光的直线方向上各放1个光电管（图6.18）我们会看到什么呢？如果情况仅仅是，光子有一半的机会走一条途径，另一半机会走另一条，那么我们应该发现其中一个检测器有一半的机会记录到光子，另一半机会是被另一个检测器记录到。然而，事情并非如此。如果两个途径的长度完全相同，则百分之百的机会是光子抵达放在原先光子运动的方向上的检测器A，而百分之零的概率是光子抵达另一检测器B—光子肯定打到检测器A上去！（正如在双缝实验中那样，我们可用上面的螺旋描述来看到这些。）

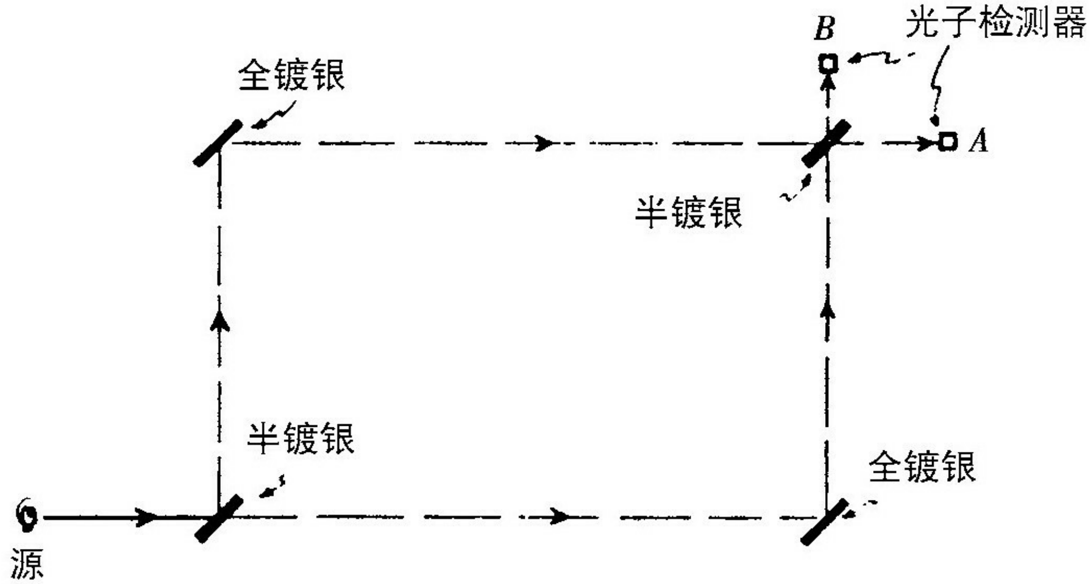


图6.18　双峰波函数的两个峰不能被简单地认为是光子在这一位置或那一位置的加权概率。可使光子所采取的两个途径相互干涉

当然，这类实验从未在路径长度达到光年数量级以上被实现过，但所叙述的结果从未被（传统的量子物理学家）认真地怀疑过！实际上，这类实验在途径长度为几米的情形下被实现过，其结果的确和量子力学的预言相一致（参阅*Wheeler 1983* ）。关于光子在它第一次和最后一次和半反射镜遭遇之间的存在的态的“实在”，此结果告诉了我们什么呢？似乎不可避免的是，在某种意义上光子实际上同时沿着两条途径旅行！因为如果将一吸收屏幕放在任何一条途径上，则光子到达A和B的概率将相等；但是当两条途径同时打开（并具有一样长度）则只能到达A！而堵住一条途径时却实际允许到达B！两条途径都打开时，光子“知道”不允许它到达B，所以它必须感觉到两条途径。

尼尔斯·玻尔关于在测量瞬息之间的光子存在没有客观“意义”的观点，依我看来是有关光子态实在的过于悲观的观点。量子力学让我们以波函数来描述光子位置的“实在”，而在半镀银镜子之间的光子波函数刚好是双峰态，双峰之间的距离有时非常可观。

我们还注意到，“同时处于两个指定的位置”不是光子态的完全描述：譬如讲我们要求能把态$\Psi_t+\Psi_b$ 从态$\Psi_t-\Psi_b$ （或$\Psi_t+i\Psi_b$ ）区别开来，这儿Ψ_t 和Ψ_b 是指分别处于两条途径中的光子（现在分别为“穿透

的”和“反弹的”光子）。正是这种区别决定了光子到达半镀银镜子时，肯定到达A或B（或以中等的概率到达A或B）。

量子实在的令人困惑的特征，也就是我们必须认真地认为的粒子可以各种（不同！）的形式“同时处于两处”——这是因为必须允许用复数权重把量子态加起来以得到其他量子态这个事实引起的。这种态的叠加是量子力学称之为量子线性叠加的一般的、重要的特征。正是它允许我们从位置态组成动量态，或从动量态组成位置态。在这些情形下，线性叠加被应用到无限多的不同的态，也就是所有不同的位置态，或所有不同的动量态。但是，正如我们已经看到的，只要把它仅仅应用于一对态就引起了这样的困惑不解。其规则是不管任何两个态是多么不同，它们能在任何复线性叠加上共存。的确，任何自身由单独粒子构成的物理对象应当能以这种在空间中分隔得很开的态的叠加的形式而存在，并因此“同时处于两处”！量子力学的形式在这方面对于单独粒子还是许多粒子的复杂系统并没有差别。那么为何我们从未经验过宏观物体，（譬如棒球，甚至人）同时处于完全不同的地方？这是一个根本的问题，今日量子理论尚不能为我们真正地提供一个满意的答案。对于像棒球这样的如此富有内容的对象，我们必须认为这些系统处于“经典水平”——或者，正如通常说的，“观察”或“测量”将对该棒球进行的——而那时对我们的线性叠加进行加权的复概率幅必须已被平方求模，并当作描述实际不同选择的概率。然而，这正好引起一个争议性问题：为何允许我们以这种方式改变U到R的量子规则！以后我还要讨论这个问题。

希尔伯特空间

我们记得在第5章为了描述经典系统引进了相空间的概念。相空间中的单独的点代表整个物理系统的（经典的）态。在量子力学中，其相应的类似概念是希尔伯特空间 [48] 。现在希尔伯特空间中的单独的点代表整个系统的量子态。我们需要浏览一下希尔伯特空间的数学结构。我希望读者对此无所畏惧，我应该说，虽然其中的一些思想也许是非常陌生的，它不是数学上非常复杂的东西。

希尔伯特空间的最基本的性质在于它是一种所谓的矢量空间——事实上，是一个复的矢量空间。这表明允许我们把空间的任何两个元素加起来得到另一个元素，也允许我们实行带有复杂权重的加法。因为这些

是我们刚刚考虑的量子线性叠加的运算，也就是对于上面光子给予我们 $\Psi_t+\Psi_b$ ，$\Psi_t-\Psi_b$ ，$\Psi_t+i\Psi_b$ 等的各种运算。我们能做到这些。我们使用的术语“复矢量空间”的所有含义就是允许进行这类带权的求和[5]。

可以十分方便地使用狄拉克引进的记号，用某种角括号诸如|Ψ〉，|x〉，|Ψ〉，|1〉，|2〉，|3〉，|n〉，|↑〉，|↓，〉，|→〉，|→〉等表示被当作态矢量的希尔伯特空间元素。这样，这些符号现在表示量子态。我们把两个态矢量的叠加写作：

|Ψ〉+|x〉，

而带复数权重w, z的求和写作：

w|Ψ〉+z|x〉

（这里w|Ψ〉表示w×|Ψ〉等。）相应地，我们现在可以将上述的组合 $\Psi_t+\Psi_b$ ，$\Psi_t-\Psi_b$ ，$\Psi_t+i\Psi_b$ 分别写为|Ψ_t〉+|Ψ_b〉，|Ψ_t〉-|Ψ_b〉，|Ψ_t，〉+i|Ψ_b〉。我们还可以将一个单独态|Ψ〉乘上一个复数w得到：

w|Ψ〉。

（这是前面的一个特例，即z=0。）

我们知道可以允许进行复权重的组合，这里w和z不必要是真正的概率幅，只要是和这些幅度成比例即可。相应地，我们采用允许以一个非零复数去乘整个态矢量而物理态不变的规则。（这会改变w和z的实际的值，但是w：z保持不变。）下面的每一矢量|Ψ〉，2|Ψ〉，-|Ψ〉，i|Ψ〉，$\sqrt{2}$|Ψ〉，π|Ψ〉，（1-3i）|Ψ〉等，正如z|Ψ〉一样，代表同一个物理态（z≠0）。希尔伯特空间唯一不能解释为物理态的要素是零矢量。（亦即希尔伯特空间的原点。）

为了对所有的这一切进行几何描述，让我们首先考虑“实”矢量的更通常的概念。人们通常将这样的矢量简单地摹想成平面上或三维空间上的一个箭头。利用平行四边形定律可得到两个箭头的和（图6.19）。用一个（实）数乘一个矢量的运算，按照“箭头”的图像就是简单地将此箭头的长度乘上这数，同时保持箭头的方向不变。如果乘数为负的，那么

箭头的方向倒过来；如果乘数为零，则得到零矢量，它没有方向。（矢量O表示零长度的“零箭头”。）作用到一个粒子上的力即是这种矢量的一个例子。而经典速度、加速度和动量则为另外的例子。还有我们在上一章结尾处考虑的动量四矢量那是在四维而不是二维或三维空间的矢量。然而，希尔伯特空间中的矢量具有更高维数（事实上，通常是无限维的，但这一点在这里并不是重要的）。我们记得在经典相空间中也用箭头来表示矢量——那一定是非常高维的。相空间的“维数”不代表通常的空间的方向，希尔伯特空间的“维数”也是这样。相反地，每一希尔伯特空间的维数对应于量子系统的不同的独立的物理态。

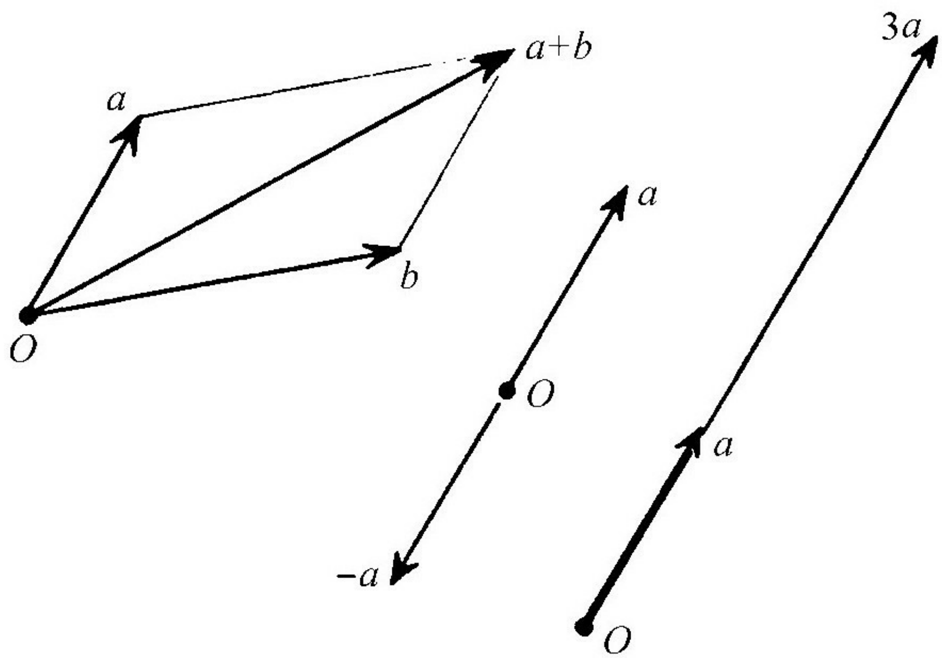


图6.19　在希尔伯特空间中的矢量加法和矢量乘以标量，可以用通常的方式，正如对在平常空间中的矢量那样摹想

由于|Ψ〉 和z|Ψ〉 是等效的，所以一个物理态实际上对应于希尔伯特空间中通过原点的整条直线或射线（表述成某一矢量的所有的倍数），而不是这条线上的某一特殊的矢量。这射线包含特定态矢量|Ψ〉的所有可能的倍数（请记住，这些是复的倍数，所以直线实际上是复的线，但是现在最好不去忧虑它！）（图6.20）。我们将很快找到二维希尔伯特空间情形下的射线空间的精巧图画。无限维的希尔伯特空间

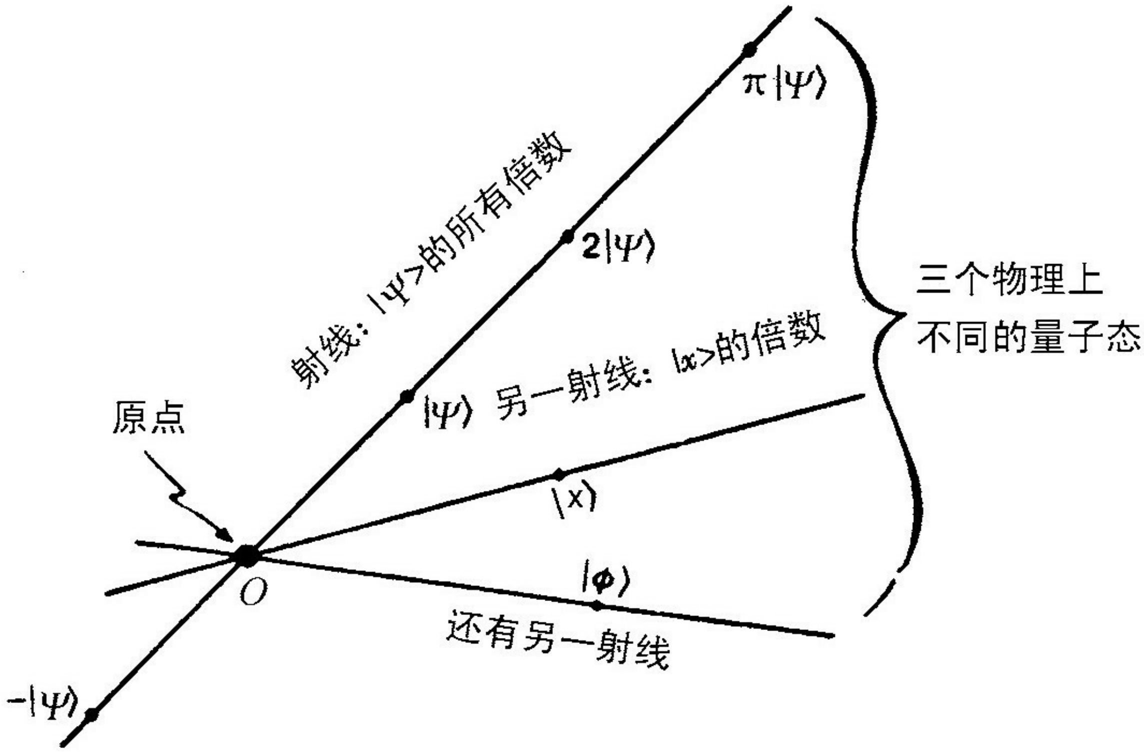


图6.20　希尔伯特空间中的整射线代表物理量子态

是另一种极端情形。甚至在简单的单独粒子位置的情形下也会出现无限维的希尔伯特空间。粒子所有可能的位置都有完整的维！粒子的每个位置都在希尔伯特空间中定义一个完整的“坐标轴”。这样，对应于粒子的无限不同的位置在希尔伯特空间中就有无限多不同的独立的方向（或“维数”）。动量态也可在同一希尔伯特空间中被表述。动量态可表达成位置态的组合，每一动量态对应于一个“对角线”出发的相对于位置轴倾斜的轴。所有动量态的集合提供了新的轴的集合。而从位置态轴向动量态轴的过渡牵涉到希尔伯特空间中的一个旋转。

人们别想以精密的方式来摹想这一切。那是不合情理的！然而，从通常的欧几里得几何可以得到某些对我们非常有用的观念。特别是，我们直到现在考虑过的轴（所有的位置空间轴或所有的动量空间轴）都认为是相互正交的，也就是相互夹角为“直角”。射线之间的“正交性”是量子力学中的一个重要概念：正交的射线是指相互独立的态。粒子所有可能不同的位置态都相互正交，所有可能不同的动量态也是如此。但是位置态并不和动量态垂直。这种情形已在图6.21上被非常概要地表达出来。

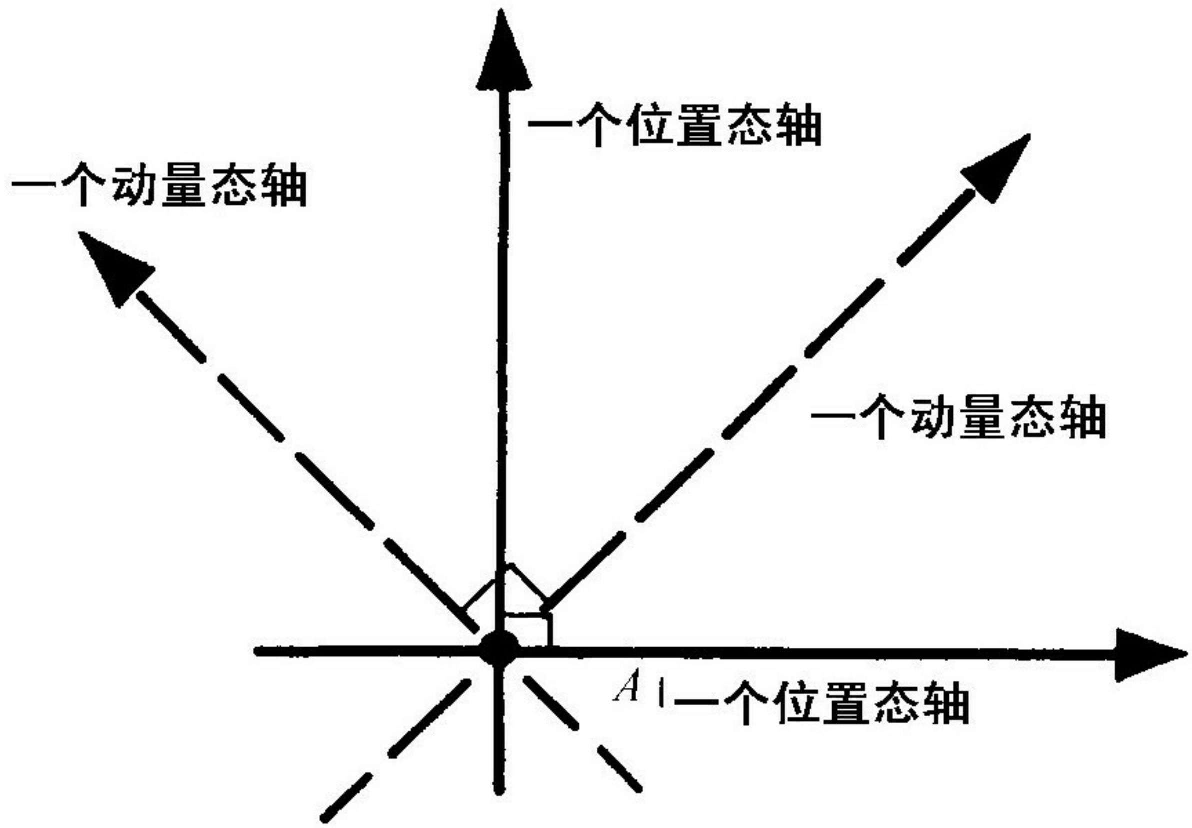


图6.21　位置态和动量态在同一个希尔伯特空间中提供了正交轴的不同选取

测量

测量（或观察）的一般规则R要求，量子系统的不同方面能被同时放大到经典水平的以及之后系统应当选取的不同状态必须永远是正交的。对于一次完整的测量，可选取的不同选择的集合组成正交基矢量的集合，表明希尔伯特空间中的每一矢量都能（唯一地）按照它们线性地表达出来。对于一个只包含单粒子的系统的位置测量，这些基矢量定义了我们刚刚考虑的位置轴。对于动量，它是定义为动量轴的不同的集合，对于不同种类完整的测量，还相应有其他的集合。测量之后，该系统的态跃迁到这些测量所决定的集合的一个轴上去——其选择只由概率来制约。没有任何动力学定律能告诉我们大自然会在已挑出的轴中选择哪一个。其选择是随机的，其概率为概率幅的平方模。

假定我们对一个具有态$|\Psi\rangle$的系统进行了完整的测量，所选择的测量的基为：

|0〉，|1〉，|2〉，|3〉，……

由于它们组成了完全集，任何态矢量，特别是|Ψ〉可以按照它们而线性地[49]表示为：

$|\Psi\rangle = z_0|0\rangle + z_1|1\rangle + z_2|2\rangle + z_3|3\rangle + \ldots\ldots$

在几何上，分量z_0，z_1，z_2，……是矢量|Ψ〉的在不同的轴|0〉，|1〉，|2〉，……上的正交投影的大小的测度（图6.22）。

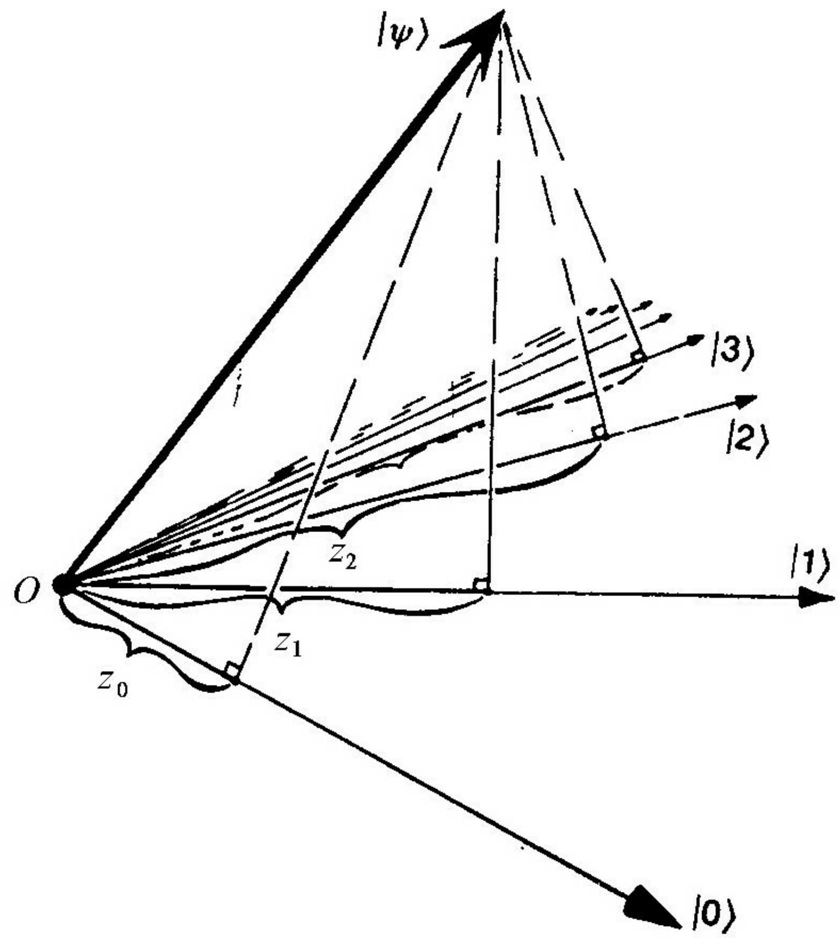


图6.22　态|Ψ〉在轴|0〉，|1〉，|2〉，……上的正交投影的大小提供了所需要的幅度z_0，z_1，z_2，……

我们能将复数z_0 ，z_1 ，z_2 ，……解释作所需要的概率幅，这样它们的平方模就提供了在测量后该系统处于相应的|0〉，|1〉，|2〉，……等态的不同概率。然而，这还不完全，因为我们还未固定住不同的基矢量|0〉，|1〉，|2〉，……的“尺度”。为此我们必须指明它们在某一种意义上是单位矢量（亦即具有单位“长度”的矢量），用数学的术语，它们组成了所谓的正交基（相互垂直的并归一化为单位矢量）[6] 。如果|Ψ〉也被归一化成单位矢量，那么所需的相应的概率$|z_0|^2$ ，$|z_1|^2$ ，$|z_2|^2$ ，……如果|Ψ〉不是单位矢量，则这些数就分别和所需的概率幅成比例。实际的幅度就为：

$$\frac{z_0}{|\Psi|}，\frac{z_1}{|\Psi|}，\frac{z_2}{|\Psi|}，等等，$$

并且实际概率为：

$$\frac{|z_0|^2}{|\Psi|^2}，\frac{|z_1|^2}{|\Psi|^2}，\frac{|z_2|^2}{|\Psi|^2}，等等，$$

这里|Ψ|是态矢量|Ψ〉的“长度”。每一态矢量都具有正实数的“长度”（除了O具有零长度），而且如果|Ψ〉为单位矢量则|Ψ|=1。

完整测量是一种非常理想的测量。例如，一个粒子的位置的完整测量需要我们能在宇宙中的任何地方以无限精度将该粒子定位！一种更初等的测量是我们简单地问是或非的问题，譬如：“该粒子是处于某一根直线的左边或右边？”或“该粒子的动量是在某一个范围内吗？”等等。是或非的测量真正是测量的最基本类型。（例如，人们可以只用是或非测量把粒子的位置或动量收缩到任意小的范围。）假定是或非测量的结果为是，那态矢量必须在希尔伯特空间的“是”的我称之为Y的区域内。另一方面，如果测量的结果为非，那态矢量就在希尔伯特空间的“非”的我称之为N的区域内。区域Y和N是完全相互正交的，任何属于Y的态矢量必须和属于N的任何矢量正交（反之亦然）。此外，任一态矢量都能以唯一的方式表达成分别来自Y和N的两个矢量之和。用数学的语言讲Y和N是相互正交互补的。这样，|Ψ|可唯一地表达成：

$|\Psi\rangle = |\Psi_Y\rangle + |\Psi_N\rangle$，

这里$|\Psi_Y\rangle$属于Y，而$|\Psi_N\rangle$属于N。$|\Psi_Y\rangle$称为态|Ψ〉在Y的正交投影。相应地，$|\Psi_N\rangle$为|Ψ〉在N上的正交投影（图6.23）。

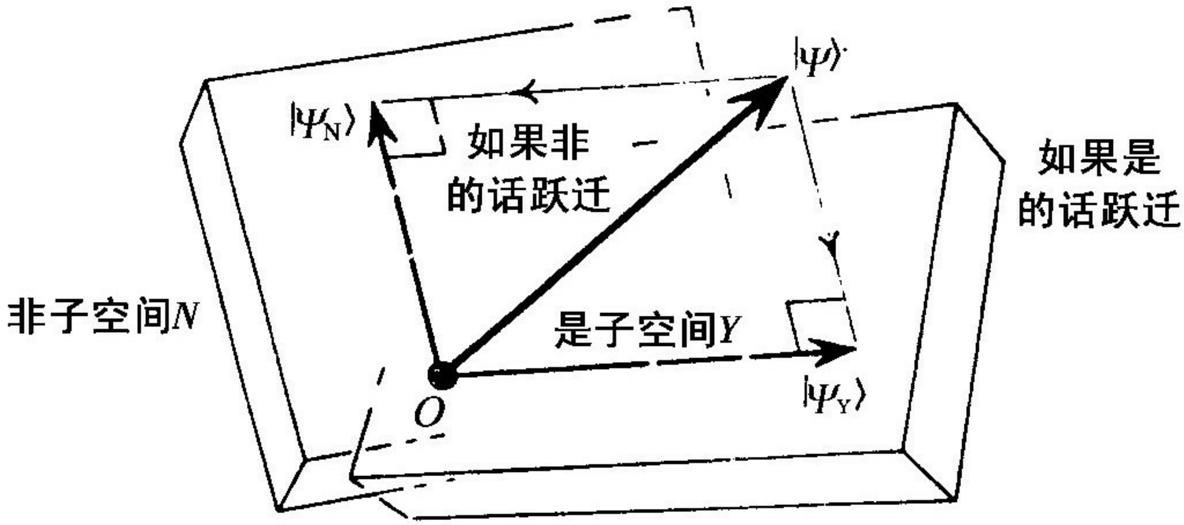


图6.23　态矢量的约化。可以按照一对相互正交互补的子空间Y和N来描述是或非测量。测量后，态|Ψ〉跃迁到它在其中一个子空间的投影，而态矢量长度平方在投影中减少的因子给出跃迁概率

在测量时，态|Ψ〉跃迁并成为（比例于）$|\Psi_Y\rangle$ 或$|\Psi_N\rangle$ 。如果结果为是，则它跃迁到|ΨY〉；如果为非，则跃迁到$|\Psi_N\rangle$ 。如果|Ψ〉是归一化的，则发生这些的相应概率为这些投影的态的长度平方：

$$|\Psi_Y|^2,\ |\Psi_N|^2。$$

如果|Ψ〉不是归一化的，我们必须将这些表示式除以$|\Psi|^2$ 。（“勾股定理”，$|\Psi|^2=|\Psi_Y|^2+|\Psi_N|^2$ 断言，这些概率之和为1，正如所预想的那样！）请注意，从|Ψ〉跃迁到|ΨY〉的概率由在投影中的长度平方的减少的比所给出。

关于作用于量子系统的“测量动作”还有最后一点要弄清。不管对于任何态——譬如态|x〉—总存在一个可在原则上进行的是或非测量[7] 。如果被测量的态是（比例于）|x〉，其答案则为是；如果垂直于|x〉则为非。这样上面的区域Y可包含任何选定的态所有的倍数。这似乎隐含有很强的意义，态矢量必须是客观存在的。不管物理系统的态是什么，我们可称之为|x〉。存在一种原则上可实行的测量，在此测量下|x〉为唯一的（只差一个比例系数）肯定得到是的结果的态。这种测量对于某

些态|x〉也许是极其困难，甚至在实际中是“不可能”实现的。但是，根据这个理论，这样的测量在原则上能实现的事实，将会在本章后面产生某些惊人的推论。

自旋和态的黎曼球面

量子力学中称为“自旋”的量有时被认为是所有物理量中最“量子力学”的。这样，我们对之稍微多加注意是明智的。什么是自旋？它本质上是粒子旋转的度量。“自旋”这个术语暗示某种像板球或棒球自旋的东西。让我们回忆一下角动量的概念，正如能量和动量一样，它是守恒的（见第5章215页和293页）。只要物体不受摩擦力或其他力的干扰，它的角动量就不随时间改变。量子力学的自旋的确是如此，但是我们这里开心的是单独粒子的“自旋”，而不是大量的单独粒子围绕着它们共同质心的轨道运动（这正是板球的情形）。物理学的一个显著事实是，自然中发现的大多数粒子在这种意义下的确是在“自旋”，每种粒子都有自己固有的自旋的大小[8]。然而，正如下面要看到的，单独量子力学粒子的自旋有一种我们绝不能从自旋着的板球等的经验所能预料到的某种特殊的性质。

首先，对于一特殊类型的粒子，其自旋的大小总是一样的。只有自旋的轴的方向可以（以一种我们就要讲到的非常奇怪的方式）改变。这和板球的情形形成全然的对比，板球可依出球方式的不同具有任意大小任意方向的自旋！对于电子、质子或中子，自旋大小总为$\hbar/2$，刚好是玻尔原先允许的一个原子的量子化的角动量的最小正值的一半。（我们记得这些值为$0, \hbar, 2\hbar, 3\hbar,$ ……）我们在这里需要基本单位的一半——而在某种意义上，$\hbar/2$ 本身是更基本的单位。只包括一些公转的粒子，而每一个粒子都不自旋的对象不允许有这个角动量值。它只能是由自旋为粒子自身的固有的性质而引起的（也就是说，不是因为它的“部分”围绕某种中心的公转引起的）。

具有自旋为$\hbar/2$ 的奇数倍（如$\hbar/2, 3\hbar/2$或$5\hbar/2$ 等）的粒子称为费米子。它在量子力学描述中呈现出非常奇怪的行径：完整的360°的旋转使态矢量回到负的态矢量，而不是回归到自身！自然界的许多粒子的确是费米子。它们古怪的形式，对我们自身的存在是如此之关键——我们在后面还要讲到。余下的自旋为$\hbar/2$ 的偶数倍，也就是是$\hbar$ 的整数倍（即

0，ℏ，2ℏ，3ℏ， ……）的粒子称作玻色子。在360°的旋转下，玻色子的态矢量回归到自身，而不是它的负矢量。

考虑一个半自旋也就是自旋值为ℏ/2 的粒子。为了确定起见，假定此粒子为电子，但质子、中子甚至某种原子的情形也是一样的。（一个“粒子”可以允许具有个别部分，只要它整个可以用量子力学处理，并具有定义得很好的角动量就可以了。）我们使电子处于静止状态，并只考虑其自旋态。现在量子态空间（希尔伯特空间）只有二维，所以我们可以采用只有两种状态的基。我把这些态标成|↑〉和|↓〉。其中|↑〉表示按右手定则垂直向上的自旋，|↓〉表示向下的自旋（图6.24）。态|↑〉和态|↓〉是相互正交的，我们并将它们归一化（$|\uparrow|^2=|\downarrow|^2=1$）。电子任何可能的自旋态都是这仅有的两个正交态|↑〉和|↓〉也就是向上和向下的态的线性叠加，譬如w|↑〉+z|↓〉。



图6.24　电子自旋态的基由两种状态组成。它们可取作自旋向上和自旋向下的两种态

关于“向上”和“向下”的方向并没有什么特别之处。我们可以一样便利地选择在任何其他方向的自旋，譬如向右|→〉和相反的向左|←〉的态去描述。然而，对于|↑〉和|↓〉的适当的复数比例的选取，我们发现[50]：

|→〉=|↑〉+|↓〉以及|←〉=|↑〉一|↓〉。

这为我们提供了新的视角：任何电子的自旋态都是两正交态|→〉和|←〉也就是向右的和向左的态的线性叠加。我们可以另外选择完全任意的方向，譬如态矢量|↗〉指定的方向。这又是|↑〉和|↓〉的某种复线性叠加，譬如：

$$|\nearrow\rangle = w|\uparrow\rangle + z|\downarrow\rangle,$$

而每一个自旋态为此态和与它正交的态$|\swarrow\rangle$指向和$|\nearrow\rangle$相反[9]）的线性叠加。（注意，在希尔伯特空间中的“正交”的概念不需要对应于通常空间的“直角”。此处正交的希尔伯特空间矢量对应于空间的相反方向，而不是两个方向成直角。）

什么是$|\nearrow\rangle$在空间中所决定的方向和两个复数w和z的几何关系呢？由于$|\nearrow\rangle$给出的物理态并不因为被用任何非零复数去乘它而改变，所以只有z和w的比才有意义。将这个比写作：

q=z/w.

q只是某个复数，除了为了和w=0的情形相一致而“q=∞”，也就是当自旋方向垂直向下也是允许的以外。除了q=∞以外，我们总能用q代表复平面上的一点，正如我们在第3章所做的。我们可以想象复平面水平地处于空间中，按上面的描述实轴的方向“向右”（亦即在自旋态|→〉的方向上）。想象一个中心在复平面原点上的单位球面，这样点1，i，-1，-i都在球面的赤道上。我们将南极上的点设定为∞，然后从该点开始投影，这样整个复平面都被映射到球面上。任何复平面上的点q都对应于球面上唯一的一点q，它可通过作过复平面上的q点与南极点这两点的直线而得到（图6.25）。这种对应称为球极平面投影。它具有美丽的几何性质（亦即它可将圆域保角映射成圆域）。该投影使我们可用复数和∞一起，也就是可能存在的复比q的集合，来标记球面上的每一点。以这种特殊方式标记的球面称作黎曼球面。

黎曼球面对于电子自旋态的意义在于，态$|\nearrow\rangle = w|\uparrow\rangle + z|\downarrow\rangle$ 的自旋方向和由从中心到黎曼球面上标记有q=z/w点的实际方向一致。我们注意到，北极对应于态|↑〉，它是z=0，也就是标记作q=0，而南极为|↑〉，标记作w=0亦即q=∞。最右的点标记着q=1，它提供|→〉=|↑〉+|↓〉，而最左的点q=-1提供了|←〉=|↑〉-|↓〉。绕过球面最远的点标作q=i，相应于态|↑〉+i|↓〉，其自旋的方向直接离开我们，而最近的点为q=-i，对应于|↑〉-i|↓〉，其自旋直接指向我们。而一般的标记为q的点对应于|↑〉+q|↓〉。

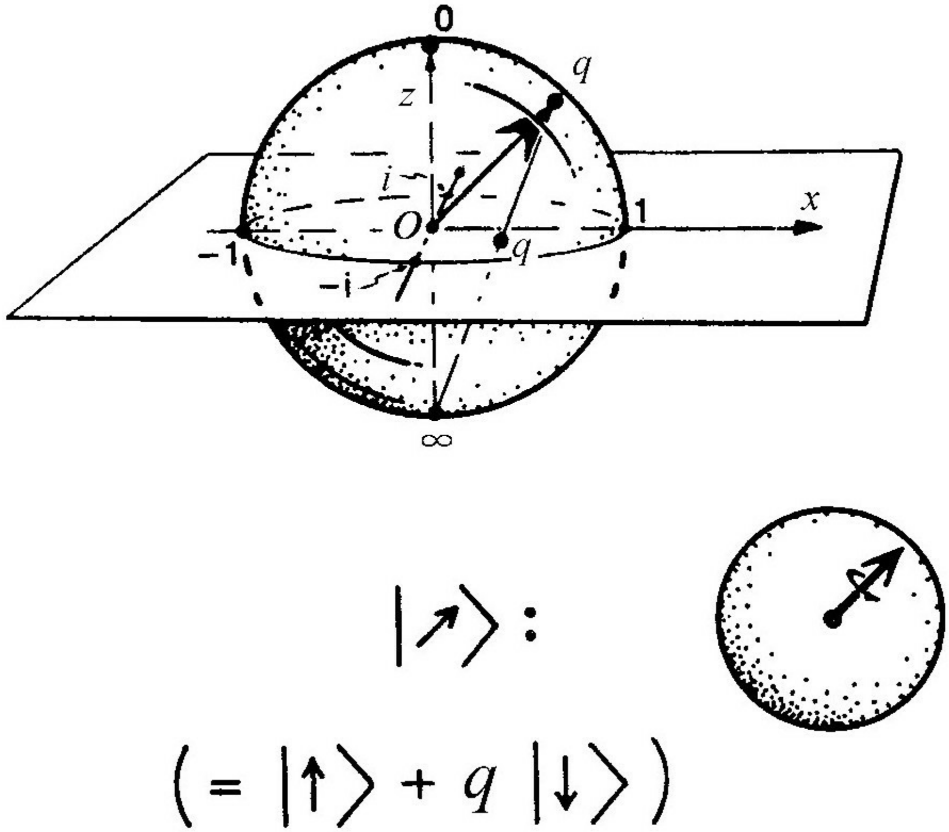


图6.25　此处用黎曼球面来表示自旋为1/2的粒子的物理上不同的自旋态。自球面南极（∞）作球极平面投影，将球面投射到通过其赤道的亚根平面上

所有这一切和人们要进行的电子自旋的测量有什么关系呢？[10] 在空间选取某一个方向；我们称为α。如果我们在此方向测量电子自旋，答案为是表明电子（现在）的确以右手定则在α方向自旋，而非表明自旋的方向和α相反。

假定答案为是，那么我们将此结果的态标记为|α〉。如果我们简单地重复此测量，利用和前面完全同样的方向α，则我们的答案应该又是百分之百的概率为是。但是如果在第二次测量时我们改变方向，改到一个新的β方向，则会发现答案为是的跃迁到态|β〉上去的概率小了。还有答案为非的跃迁到和β相反方向的态上去的概率。如何计算此概率呢？答案是在上节结尾处的方案中。第二次测量为是的概率为：

1/2（1+cosθ），

这里θ是两个方向α，β之间的夹角[11]。相应地，第二次测量为非的概率为：

1/2（1-cosθ）。

我们从这里能看到，如果第二次测量是在与第一次夹角直角的情况，则两种结果的概率都为百分之五十（cos90°=0）：第二次测量的结果完全是随机的！如果两次测量的夹角为锐角，则答案为是的可能性比非要更大。如果为钝角，则非的可能性更大。在β和α相反的极端情形下，答案为是的概率为0，而为非的概率为百分之百；也就是说，第二次测量的结果一定是和第一次相反。（参见*Feynman et al.1965* 关于自旋的更详尽的讨论。）

黎曼球面实际上对于任何双态的量子系统，在描述一系列可能的量子态（准确到一个比例系数）时起着基本的（但是未被广泛认识到的）作用。对于半自旋的粒子，它的几何作用特别明显，因为球面上的点对应于自旋轴的可能的空间方向。在其他很多情形，难以看到黎曼球面的作用。考虑刚刚通过双缝隙，或从半镀银镜子反射回来的光子。光子态为某个描述两个完全不同位置的双态$|\Psi_t\rangle$和$|\Psi_b\rangle$的诸如$|\Psi_t\rangle+|\Psi_b\rangle$，$|\Psi_t\rangle-|\Psi_b\rangle$或$|\Psi_t\rangle+i|\Psi_b\rangle$等的线性组合。黎曼球面仍然描述物理上一系列不同的可能性，但现在仅仅是抽象地。态$|\Psi_t\rangle$由北极（“顶”），$|\Psi_b\rangle$由南极（“底”）分别代表。而$|\Psi_t\rangle+|\Psi_b\rangle$，$|\Psi_t\rangle-|\Psi_b\rangle$以及$|\Psi_t\rangle+i|\Psi_b\rangle$由赤道上的不同的点代表。一般地，$w|\Psi_t\rangle+z|\Psi_b\rangle$为点q=z/w所代表。在很多情况下，正像这个例子，“黎曼球面可能的价值”相当隐蔽，和空间几何没有明显的关系。

客观性和量子态的可测量性

尽管我们在正常的情况下只能为实验的结果提供概率的这个事实，关于量子力学的态似乎有某些客观的东西。人们经常断言，态矢量只为了方便描述“我们已知”的物理系统——或者，态矢量也许实际上并不描述一个单独的系统，而仅仅是提供大量制备好的类似系统在“系综”方面的概率信息。在关于量子力学告诉我们物理世界的实在性方面，我觉得这种意见过分胆怯。

有关态矢量的“物理实在性”的一些谨慎或怀疑，是由于按照该理论，物理上可测量的东西严格受到限制这个事实引起的。让我们考虑上述的电子自旋态。假定自旋态刚好是|α〉，但是我们不知道这些，也就是说我们不知道电子自旋的方向α。我们能否用测量来决定此方向呢？不，我们不能。我们最多能做的只是提取“部分”信息——就是简单的是或非问题的答案。我们可以选取空间中的某个方向β并在该方向上测量电子自旋。我们得到的答案非是即非，但在此之后，我们就丧失了关于原先自旋方向的信息。答案为是的话，我们知道现在这个态和|β〉成比例；答案为非的话，则现在的态在和β相反的方向上。没有任何一种情形告诉我们测量之前态的方向α，它仅仅是给出了关于α的某种概率的信息。

另一方面，似乎有某种完全客观的关于方向α的东西，电子在测量之前“刚好沿着这个方向自旋”[51]。由于我们也许选定了在方向α上测量电子的自旋——而电子必须肯定地给出的答案，如果我们刚好猜中了的话！无论如何，电子的自旋态中储藏着电子实际上必须给出的这个答案的“信息”。

我似乎觉得，在按照量子力学来讨论物理实在的问题时，我们应该将什么是“客观的”和什么是“可测量的”区别开来。在对一个系统进行实验时，不能准确地（除了比例系数外）断定它处于何态，也就是说系统的态矢量的确是不可测量的。但是，态矢量似乎的确（又是除了比例系数外）是系统的完全客观的性质，它为人们可能进行的实验的结果所完全表征。在诸如电子的半自旋的单独粒子的情形，因为它仅仅断言存在电子自旋被精确定义的某方向，即便也许我们不知道这个方向，这种客观性也不是不合理的。（然而，以后我们会看到，对于更复杂的系统，这个“客观的”图像会变得更奇怪得多——甚至对于仅仅包含一对半自旋粒子的系统而言也是如此。）

但是，在电子自旋被测量之前它必须有一个物理上定义的态吗？在许多情形下，它没有必要。因为它自身不能被认为是一个量子系统，物理态一般地必须认为是一个和其他大量粒子纠缠在一起的电子的描述。然而在特殊情形下，可以考虑电子本身（至少就其自旋而言）。按照标准的量子理论，在这种情况下，譬如它的自旋的方向预先（也许未知的时刻）被测量过之后的一段时间内没受到干扰，那么电子就具有完全客观的定义好的自旋方向。

复制量子态

电子自旋态的客观性以及不可测量性阐释了另一个重要事实：不能在使原先的态不被触动的情形下将其复制。因为假定我们能对一个电子的自旋态|α〉进行复制。若能复制一遍，则能两遍多遍地复制。结果的系统会在一个定义得非常好的方向上具有大的角动量。可由宏观测量把这个方向|α〉确定下来。这就违反了自旋态|α〉的基本的不可测量性。

然而，如果我们准备去破坏原先的态，则复制便成为可能。例如，我们有一处于未知的自旋态|α〉的电子和另一处于另一个自旋态|γ〉的中子。将它们交换，使中子自旋态为|α〉而电子态为|γ〉是完全合法的。我们所不能做的是复制|α〉。（除非我们预先知道|α〉实际上为何态！）（还可参阅*Wootters and Zurek 1982*）

我们记得在第1章（31页）讨论过"远距运送机器"。这机器，原则上依赖于在遥远的行星上有可能拼装出一个人的身体大脑的复制本。一个人的"所知所闻"可以依赖于一个量子态的某些方面，这是一个令人感兴趣的猜想。若果真如此，则量子力学禁止我们去复制"所知所闻"而不破坏原先的态。远距离搬运的"矛盾"可望以这种方式得到解决。量子效应和大脑功能的可能关联将在最后两章考虑。

光子自旋

让我们在下面考虑光子的"自旋"以及它和黎曼球面的关系。光子具有自旋，但是因为它们总是以光速运动，人们不能将自旋认为是围绕于一个固定点；相反的自旋轴总在运动的方向。光子自旋称之为极化，这就是"偏振片"太阳镜的行为所根据的现象。把两块偏振片重叠在一起并透视之。一般地讲，你会发现有一定量的光透过去。现在使其中一块不动而旋转另一块，通过的光量会发生变化。在一个方向上，穿透的光达到最大，第二块偏振片实际上并没减少穿透的光量；在与此垂直的方向上，第二块偏振片可使通过的光量减少到零。

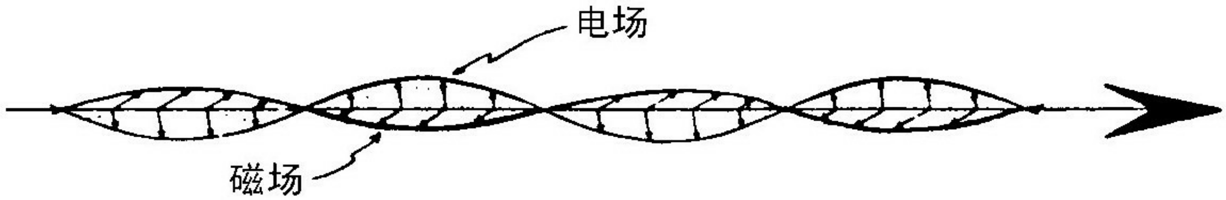


图6.26　平面偏振的电磁波

按照光的波动图像最容易理解所发生的现象。在这里我们需要用麦克斯韦的光波的振动电磁场描述。图6.26画出了平面偏振的光。电场在一个称为极化面的平面上上下振动。而磁场在一个垂直于电场振动的平面上振动，电磁场相互共振。每一块偏振片让极化面和偏振片结构相平行的光通过。当第二块偏振片的结构和第一块指向一致时，所有通过第一块偏振片的光就会通过第二块偏振片。但是，当它们结构的方向相互垂直时，第二块偏振片就将通过第一块偏振片的光全部阻拦住。如果两块偏振片的指向夹角为φ时，则第二块偏振片让

$$\cos^2 \varphi$$

部分的光通过。

在粒子表像中，我们应该把每一单独光子认为是具有偏振的。第一块偏振片的行为像一个偏振度测量器。如果光子的确在一个合适的方向偏振，它就给出是的答案，并让光子通过。如果光子在与此相垂直的方向偏振，则答案为非，光子就被吸收。（注意在希尔伯特空间中的“正交”并不对应于通常空间中的“相交成直角”！）假定光子通过了第一块偏振片，则第二块偏振片就会问相应的问题，但是对于某个其他的方向。如果两个方向的夹角为φ，我们现在就有$\cos^2 \varphi$作为已经通过第一块偏振片的光子通过第二块偏振片的概率。

黎曼球面和这些有何相干呢？为了得到偏振态的全部复数系列，我们必须考虑圆的和椭圆的偏振。图6.27画出了经典波动的情形。圆偏振时电场旋转，而不是振荡。磁场仍然和电场成直角并同步地旋转。椭圆偏振可看成旋转和振动的结合，而描写电场的矢量在空间划出一个椭圆。在量子描述中，每一单独光子允许这些不同极化的方式——光子自旋的态。

如何在黎曼球面上将所有这些可能性表示出来呢？想象一个垂直向上运动的光子。现在北极代表右手自旋的态|R〉，这表明当光子通过时电场矢量以反时针方向绕着垂直的轴旋转（从上面看）。而南极代表左手自旋的态|L〉。（我们可以把光子想象成像来复枪子弹一样自旋，或是右旋或是左旋。）一般的自旋态|R〉+q|L〉是这两种态的复线性组合，它对应于黎曼球面上标出的一点。为了求出q和偏振椭圆的关系，我们首先取q的平方根p：

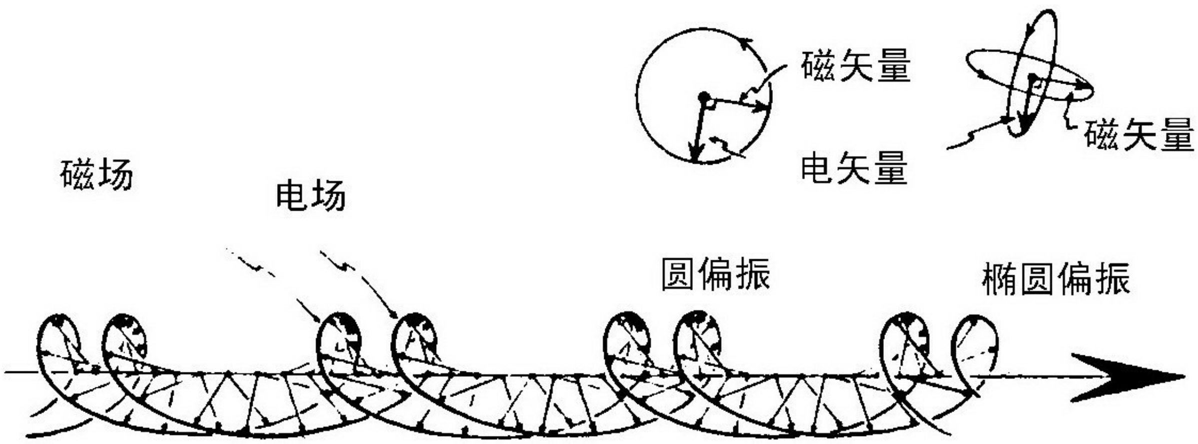


图6.27　圆偏振电磁波（椭圆偏振是介于图6.26和图6.27之间的中间情况）

$p=\sqrt{q}$。

然后在黎曼球面标出p而不是q。考虑通过球面中心的一个平面，该平面垂直于连接标上p的点和球心的直线。此平面和球面的交线为一圆周。我们将此圆周垂直投影就得了偏振椭圆（图6.28）[52]。q的黎曼球面仍然描述了光子偏振态的总体，但是q的平方根为之提供了空间实现。

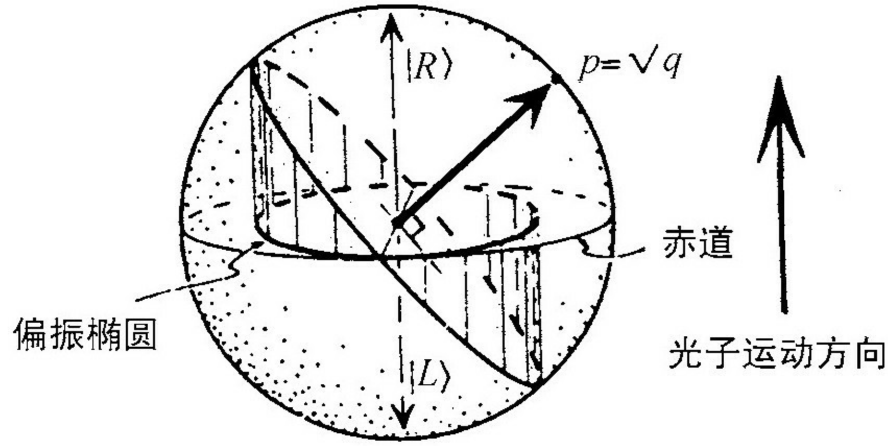


图6.28　黎曼球面（现在是√q的）也描述了一个光子的偏振态（指向√q的矢量称为斯托克斯矢量）

我们可同样地将用于电子的同一个公式1/2（1+cosθ）用于计算概率，只要我们把它应用于q而不是p。考虑一平面偏振，我们首先在一个方向上，然后在另一和它夹φ角的方向上测量光子的偏振。这两个方向对应于球面赤道上从中心看张角为φ的两个p值。因为p为q的平方根，所以q点在中心的张角为p点张角的两倍：θ=2φ。这样，在第一测量结果为是后第二测量结果亦为是（亦就是通过第一块偏振片的光子再通过第二块偏振片）的概率为1/2（1+cos2φ），这正是前面断言的$\cos^2$ φ（可用简单的三角学知识验证之）。

大自旋物体

对于具有多于两个基本态的量子系统，在物理上可区别的态的空间比黎曼球面更复杂。然而在自旋的情况，黎曼球面本身总是起着直接的几何作用。考虑以下有质量的自旋为$n\times\hbar/2$ 的粒子或原子，让它处于静止。这样自旋就定义了一个n+1态的量子系统。（对于一个无质量的，也就是以光速运动的自旋的粒子，譬如光子，正如上面所描述的，自旋总是一个两态系统。但是对于有质量的粒子，态的数目随着自旋而增加。）如果我们选择在某一个方向测量该自旋，会发现共有n+1个不同的可能的结果，此结果依自旋相对于该方向的指向而定。按照基本的单位$\hbar/2$ ，在那个方向自旋的可能结果为n, n-2，n-4，……，2-n或-n。这样n=2时其值为2，0或-2；n=3时其值为3，1，-1或-3；等等。负值对应

于自旋主要指向和所测量的方向相反的方向。在半自旋的情形，亦即n=1时，上述的值1对应于是，而值-1对应于非。

由于我不想企图在这里解释的原因，人们发现（*Majorana 1932*，*Penrose 1987a*）对于$\hbar n/2$的自旋每一个自旋态（准确到一个比例系数）可唯一地由黎曼球面上的（无序的）n点的集合，也就是从中心出发的n个（通常不同的）方向表征（图6.29）。这些方向由可能对此系统进行的测量所表征：如果我们在它们中的任一个方向测量自旋，则结果一定不会全在相反方向上，也就是给出值n, n-2，n-4，……，2-n，但不会有-n。在譬如上述电子的n=1的特殊情形下，这就是在上面描述中标以q的黎曼球面上的一点。但是对于大数值的自旋，正如我刚才描述的，图像变得更为精巧——虽然，由于某种原因，物理学家对此并不特别熟悉。

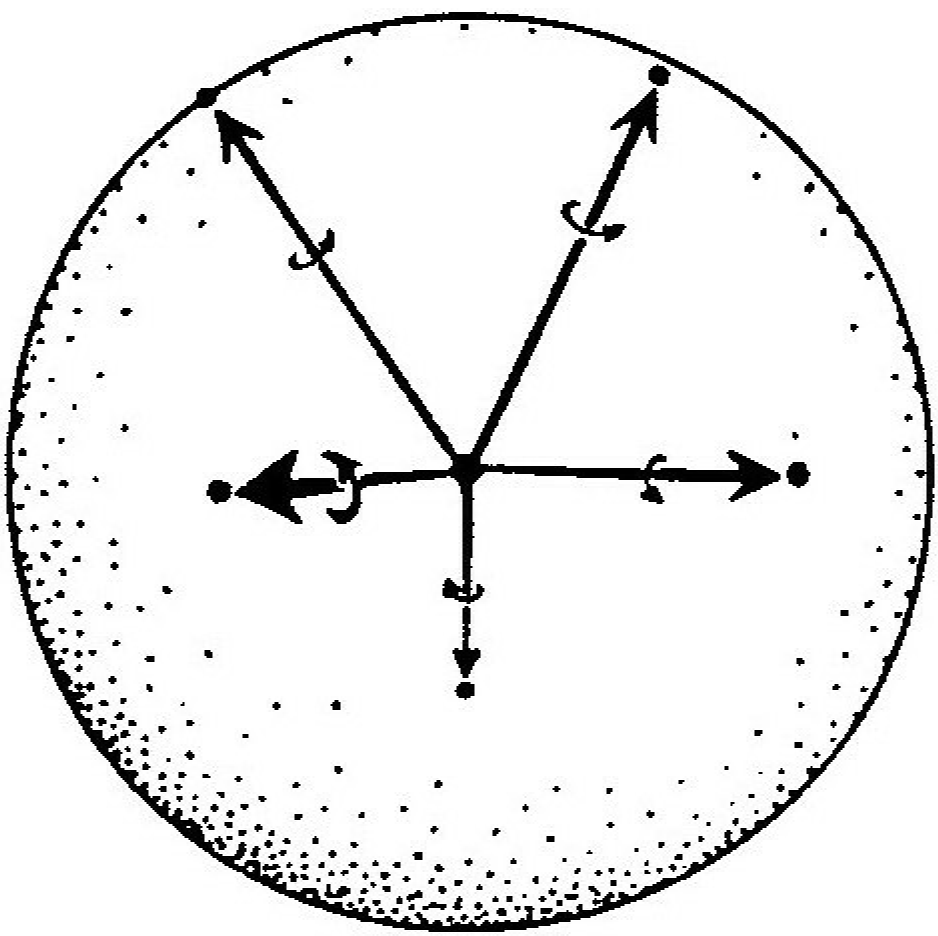


图6.29　对于一个有质量的粒子，一般的高自旋态可用指向任意方向的半自旋态的集体来描述

在这些描述中有些相当令人吃惊和困惑的东西。人们经常相信，当系统变得更大更复杂时，在某种适当的极限的意义上，原子（或基本粒子或分子）的量子描述就会过渡到经典的牛顿描述。然而，在实际情况中，这肯定是不对的。正如我们已经看到的，具有大角动量的客体的自旋态对应于大量的杂乱地撒开在黎曼球面上的点 [53] 。我们可以把物体的自旋认为是由一大堆大小为一半的，方向由这些点决定的自旋所组成。这些结合态中只有很少情形，其大部分点集中在球面上的一个小区域中（亦即大部分半自旋近似地指向同一个方向）——这些才对应于人们通常在譬如板球等经典物体处遇到的角动量的实际的态。我们也许会预料到，如果我们选择一个总角动量为某个非常大的数（按照单位$\hbar/2$），但是处于“紊乱”的自旋态，那么某种类似于经典自旋的东西就会开始出现。但是情况根本不是这样，一般地讲，具有大的总自旋的量子自旋态和经典态毫不相像！

那么经典物理中的角动量的对应物是如何构成的呢？大多数大自旋量子态实际上不和经典的东西相类似，它们是每一个都类似于经典的（正交的）态的线性叠加。对此系统进行“测量”时，其状态（以某种概率）“跃迁”到这一个或那一个类经典的态上去。这种情形和系统的任何其他经典的可测量的性质相类似，而不仅仅是角动量。正是量子力学这个方面在一旦系统“到达经典水平”时即起作用。在后面我还要仔细讨论这些，但在讨论这么“大”或这么“复杂”的量子系统之前，我们必须对量子力学如何实际处理包含多于一个粒子的系统的古怪方式有些了解。

多粒子系统

很不幸，多粒子状态的量子力学描述是相当复杂的。事实上，它们会变得极其复杂。人们必须按照所有粒子各自所有可能的不同位置的叠加来思考！这导致可能状态的极庞大的空间——比在经典理论中的一个场大得多了。我们已经知道，甚至在单粒子的量子态，也即一个波函数即有一整个经典场的复杂性。这个图像（需要无限个参数才能指明）已经比粒子的经典图像（这里只需几个参数就能指明其状态——如果没有内部自由度，譬如自旋的话，实际上是6个，参阅第5章228页）复杂得很多。这似乎很糟糕。人们也许以为，必须用两个场来描述两个粒子的量子态。根本不是这回事！两个或更多粒子的状态的描述，正如我们将看到的，要比这个更精巧得多！

一个单独的（无自旋的）粒子的量子态由粒子所能占领的每一可能位置上的一个复数（幅度）所定义。粒子在点A有一幅度，在点B有一幅度，在点C有一幅度，等等。现在考虑两个粒子。譬如，第一个粒子可能待在A，而第二个粒子待在B这种可能性必须有一幅度。另外，第一个粒子可待在B，而第二个粒子待在A，这也需要一幅度；或第一个粒子待在B，而第二个粒子待在C；或者也许两个粒子都在A。每一种可能都有一个幅度。这样，波函数不仅仅是位置的一对函数（也就是一对场）；它必须是两个位置的一个函数！

为了估计一个双位置的函数比两个单位置的函数复杂多少，我们可想象一种情景，只存在有限数目的允许位置的集合。假定只有10个允许的由（正交）态给定的位置：

$|0\rangle$ ，$|1\rangle$ ，$|2\rangle$ ，$|3\rangle$ ，$|4\rangle$ ，$|5\rangle$ ，$|6\rangle$ ，$|7\rangle$ ，$|8\rangle$ ，$|9\rangle$.

粒子态$|\Psi\rangle$为某种组合：

$|\Psi\rangle=z_0|0\rangle+z_1|1\rangle+z_2|2\rangle+z_3|3\rangle+\ldots\ldots+z_9|9\rangle$ ，

此处不同分量z_0 ，z_1 ，z_2 ，……，z_9 分别顺序地提供了粒子在每一点处的幅度。10个复数指定了粒子的状态。对于双粒子状态，我们对每一对位置都需要一个幅度。共有

$10^2=100$

个不同的（有序）位置对，所以我们需要100个复数！如果我们只有两个单粒子态（亦即“位置的两个函数”而不是上面的“一个双位置的函数”），则我们只需要20个复数。

我们可以把这100个数标为：

z_{00}，z_{01}，z_{02}，…，z_{09}，z_{10}，z_{11}，z_{12}，…z_{20}…z_{99}，

以及把相应的（正交）基矢量标为：[12]

$|0\rangle|0\rangle$ ，$|0\rangle|1\rangle$ ，$|0\rangle|2\rangle$ ，……，

|0〉|9〉，|1〉|0〉，……，|9〉|9〉，

则一般的双粒子态|Ψ〉可写成：

|Ψ〉=z_{00} |0〉|0〉+z_{01} |0〉|1〉1+……+z_{99} |9〉|9〉。

此处态的“乘积”记号具有如下意义：如果|α〉是第一个粒子可能的态（不必是位置态），而|β〉为第二个粒子的可能的态，则断言第一个粒子的态为|α〉以及第二个态为|β〉的态可写作：

|α〉|β〉。

可对任何其他的量子态而不必仅仅是单粒子态取“乘积”。这样，我们总是将乘积态|α〉|β〉（不必为单粒子的态）解释作描述以下事件的同时发生：

“第一系统处于态|α〉而且第二系统处于态|β〉。”

（可对|α〉|β〉|γ〉等进行类似的解释；见下面。）然而，一般双粒子态实际上并不具备这种“乘积”的形式。例如，它可以为：

|α〉|β〉+|ρ〉|σ〉，

此处|ρ〉为第一系统的另一个可能的态，而|σ〉是第二系统的另一个可能的态。此状态是一线性叠加；也就是第一个（|α〉以及|β〉）的同时发生加上第二个（|ρ〉以及|σ〉）的同时发生，而它不能被重写成一个简单的乘积（亦即作为两个态的同时发生）。作为另一例子，态|α〉|β〉-i|ρ〉|σ〉描述另一个不同的线性叠加。注意量子力学需要很清楚地区别“以及”和“加”这两个词。在现在语言中——譬如在保险小册子中——非常不幸地将“加”在“以及”的意义上使用。这里我们要加倍小心！

3个粒子的情形非常类似。在上述的只有10个可选择的位置的情况下，为了指明一般的三粒子状态，我们现在需要一千个复数！三粒子态的完备基是：

|0〉|0〉|0〉，|0〉|0〉|1〉，|0〉|0〉|2〉，

……，|9〉|9〉|9〉.

特殊的三粒子态具有如下形式：

|α〉|β〉|γ〉

（这里|α〉、|β〉和|γ〉不必为位置态），但是对于一般的三粒子态，人们必须将许多这种简单的“乘积”叠加起来。对于四个或更多粒子的相应的模式则不必赘述。

迄今为止我们只是讨论可辨别的粒子。这里我们将“第一个粒子”，“第二个粒子”和“第三个粒子”等都当作不同种类的。然而，量子力学的一个显著特点是，等同粒子的规则与上面不同。其规则事实上是，在很清楚的意义上，特别种类的粒子必须完全等同，而不仅仅是极端接近于等同。但是，所有电子之间相互等同的方式和所有光子的方式不同。粒子的这两种一般种类必须以相互不同的方式处理。

为了不使读者在完全被用词不当所混淆之前，让我首先解释费米态和玻色态实际上是如何表征的。其规则如下。如果|Ψ〉是牵涉到某一特别种类的一些费米子，那么如果两个费米子相互交换，则|Ψ〉必须作如下的变化：

|Ψ〉→-|Ψ〉。

如果|Ψ〉牵涉到某一特别种类的一些玻色子，则其中任何两个玻色子交换时，|Ψ〉必须作如下变化：

|Ψ〉→|Ψ〉。

它的一个含义是两个费米子不能处于同一态中。因为如果这样的话，把它们交换就根本不影响其总的态，我们就必须有-|Ψ〉=|Ψ〉，也就是|Ψ〉=0，对于量子态来说这是不允许的。这个性质称为泡利不相容原理[13]，它对物体的结构具有基本的含义。物体的主要成分的确是费米子、电子、质子和中子。若没有不相容原理，物体就会向自身坍缩！

我们来重新考虑10个位置的情形。我们假定有一个含有两个等同费米子的态。态|0〉|0〉被泡利原理所排除（在第一个因子和第二个因子

交换时它保持不变并没有反号）。而且，|0〉|1〉就这样子也是不行的，由于在交换时没有变成它的反号；但是这很容易由下式予以补救：

|0〉|1〉-|1〉|0〉

（如果需要的话，为了归一化，可以加上一个总的因子1/$\sqrt{2}$）。此态在两粒子相互交换时正确地变号。但现在|0〉|1〉和|1〉|0〉不再分别为独立的态。我们现在只许用一个态来取代这两个态。总之，共有

1/2（10×9）=45

个这类的态，每一个态是从不同的|0〉，|1〉，……，|9〉态的无序对而来。这样，需要45个复数才能指明我们系统的态。对于3个费米子，人们需要3个不同的位置，而基本的态看起来像下面的样子：

|0〉|1〉|2〉+|1〉|2〉|0〉+|2〉|0〉|1〉

-|0〉|2〉|1〉-|2〉|1〉|0〉-|1〉|0〉|2〉，

总共有（10×9×8）/6=120态，这样需要用120个复数去指明三费米子态。更多费米子的情形是类似的。

对于一对等同的玻色子，独立的基本态共有两类，即像

|0〉|1〉+|1〉|0〉

的态和像

|0〉|0〉

的态（现在这是允许的），共有（10×11）/2=55态。这样我们的双玻色子态需要55个复数。对于三玻色子共有3种类型的基本的态，共需要（10×11×12）/6=220个复数，等等。

当然，为了表达主要的观念，我在这里考虑简单化的情形。更现实的描述则需要位置态的整个连续统，但其基本思想是一样的。另一微小的复杂性是自旋的参与。一个半自旋的粒子（必须为费米子）在每一个位置都有两个可能的态。我们可以把它们标作“↑”（自旋“向上”）

和“↓”（自旋“向下”）。在我们简化的情况下，对于每一个粒子共有20个而不是10个基本的态：

|0↑〉，|0↓〉，|1↑〉，|1↓〉，|2↑〉，|2↓〉，

……，|9↑〉，|9↓〉，

但是除此以外，所有讨论都和以前一样地进行。[这样，对于两个这样子的费米子人们需要（20×19）/2=190个数；对于3个则需要（20×19×18）/6=1140个数，等等。]

我在第1章提到了这样的一个事实，根据现代理论，如果一个人的身体中的一个粒子和他的屋子的砖头中的一个粒子相交换，则根本不会有什么事会发生。如果那一个粒子为玻色子，正如我们看到的，态|Ψ〉的确完全不受影响。如果该粒子为一个费米子，则态|Ψ〉将由-|Ψ〉所替换，在物理上它和|Ψ〉是等同的。（如果我们感到有必要，可以修补这一符号改变，在交换之时简单地将粒子旋转360°就可以了，我们记得在进行360°旋转时，玻色子不受影响而费米子变号！）现代理论（大约在1926年）的确告诉我们有关物理物质的个别本体的问题的某些基础的东西。严格地讲，人们不能提到“这个特别的电子”或“那个单独光子”。断言“第一电子在这里而第二电子在那里”是声称态具有|0〉|1〉的形式。正如我们已经看到的，这对于费米子态是不允许的！然而，我们可以讲“存在一对电子，一个在这里，另一个在那里”，可以合法地说所有电子或所有质子或所有光子的集团（虽然在这里不管不同种类的粒子之间的相互作用）。许多单独电子为这个总图像提供一个近似，正如许多单独的质子或光子那样。这个近似在大多数目的下相当有效，但在其他一些情形下失效，超导、超流和激光的行为是众所周知的反例。

量子力学呈现的物理世界根本不是我们在经典物理中习惯了的图像。请赶紧抓牢你的帽子——量子世界中还有更为怪异的现象！

爱因斯坦——波多尔斯基——罗森“佯谬”

正如在本章开头提到的，阿尔伯特·爱因斯坦的观念，对于量子理论的发现是相当根本的。我们记得早在1905年，正是他曾先提出了“光子”的概念——电磁场的量子，由此而发展了波粒二象性的观念。（“玻

色子”的概念，正如许多其他的思想也是一部分属于他的，这在理论中占有中心地位。）然而，爱因斯坦从未接受后来从这些思想发展而来的这一个理论，他认为这理论只不过是物理世界的临时性描述。他对于这一个理论的概率方面的厌恶是众所周知的，这集中表现在他在1926年致马克斯·玻恩的回信之中（引用于Pais 1982，P443）：

量子力学是令人印象深刻的。但是一个来自内部的声音告诉我，它还不是事物的真谛所在。该理论虽然富于成果，但是却几乎没有在接近古老的神秘方面使我们往前迈出一步。无论如何，我坚信：不玩骰子。

然而，比这物理学的非决定论性更甚的，也是最困扰爱因斯坦的是，量子力学的描述方式明显地缺乏客观性。我在解释量子理论时竭尽全力地强调，该理论所作的世界描述，虽然经常是非常古怪和反直观的，却是真正客观的。相反地，玻尔似乎认定（在测量之间）系统的量子态并没有物理的真正的实在，只不过是关于该系统的“某人知识”的总结而已。难道不同的观察者会有关于同一个系统的不同知识，这样波函数变成某种根本上主观的——或“完全在物理学家头脑中的”某种东西？许多世纪以来，我们发展的美妙无比而精确的物理图像不应该完全消失掉；所以玻尔在经典水平上认为世界确实具有客观的实体。而似乎作为它这一切的基础的量子水平态却不具有“实在性”。

爱因斯坦完全拒绝这样的图像，他相信甚至在量子力学的微小尺度下，必须存在一个客观的物理世界。在他和玻尔之间的长期论战中，他企图（但没有成功）指出在事物的量子图像中的固有的矛盾，在量子理论之下还必须有另一个更深的结构，或许这一个结构和经典物理呈现给我们的图像更相似。也许一种我们没有直接知识的、系统的、更小的基元或“部分”的统计作用，是量子系统的概率行为的基本原因。爱因斯坦的追随者，尤其是大卫·玻姆，发展出一种“隐变量”的观点。按照这种观点，的确有某种确定的存在，但是我们不能直接得到精确定义一个系统的参量，由于在测量之前不知道这些参数值，所以产生了量子的概率。

这种隐变量理论能与量子物理所观察到的所有事实相一致吗？只要隐参数能瞬息地影响任意远的区域，也就是理论本质上是非定域的，则答案似乎是肯定的！那也不会使爱因斯坦高兴，特别是由于它引起了和狭义相对论冲突的困难。我在以后再考虑这些。最成功的隐变量理论称

为德布罗意——玻姆模型（*Broglie 1956*，*Bohm 1952* ）。由于本章的目的是对标准的量子理论，而不是对不同的竞争设想的总括，所以我不在这里讨论这些模型。如果人们需要物理的客观性，但又准备免除决定性，则标准理论本身就已足够了。人们简单地以为态矢量提供了“实在”——它通常按照平滑的决定性的步骤U演化，但是只要有效应将其放大到经典水平，它就要按照R作古怪的跃迁。然而，非定域性和相对论的明显困难依然存在。让我们浏览一下这些问题。

假定我们有一个包含两个子系统A和B的物理系统。例如，A和B可以是两个不同的粒子。假定A的状态有两个（正交的）选择|α〉和|ρ〉，而状态B可为|β〉和|σ〉。正如上面看到的，一般的结合态不是简单地为A的一个态和B的一个态的积（“并且”），而是这种乘积的叠加（“加”）。（我们说A和B是相关的。）让我们假定此系统的态为：

|α〉|β〉+|ρ〉|σ〉。

现在对A进行一个是或非的测量，将|α〉（是）从|ρ〉（非）中辨别出来。B发生了什么呢？如果测量的结果为是，那结果的态应为：

|α〉|β〉，

而如果结果为非，则结果的态是：

|ρ〉|σ〉。

这样我们测量A会引起状态B的跃迁：在答案为是时它跃迁到|β〉，而在答案为非时跃迁到|σ〉！粒子B根本没必要处在靠近A的任何地方；它们可以相距1光年那么远。然而，B的跃迁和A的测量是同时发生的！

但是，且慢！读者会说，这些被断定为“跃迁”的究竟是怎么回事？为何事情不像下面所描述的那样呢？想象一个盒子并事先知道里面装有一个黑球一个白球。假定取出这些球，把它们放在屋子的两个相反的角落里，并且没有一个球被看到。然后审视其中一个球并发现是白的（正如上述的|α〉）——嘿，奇怪！另一球变成黑的（如同|β〉）！如果发现第一球是黑的（|ρ〉），则一眨眼间第二球的不确定态就跃迁到“肯定是白的状态”（|σ〉）。读者会坚持道，没人在他或她头脑中会把第二球从“非确定的”状态到“肯定是黑的”或“肯定是白的”的突变归结为某种神

秘的非定域性的从考察第一球的时刻瞬息间传来的“影响”。

但是，自然界实际上比这更不寻常得多。在上述实验中，我们的确可以想象在测量A之前系统已经“知道”，譬如讲B的状态为|β〉而A的状态为|α〉（或B是|σ〉而A是|ρ〉），只不过实验者不知道而已。在发现A是|α〉后，他简单地推断B应处于|β〉。这是一种“经典的”观点——正如在定域的隐变量理论中一样——在实际上并没有发生物理的“跃迁”（所有都是在实验者的头脑中进行的！）根据这样的一种观点，系统的每一部分在事先“知道”任何要对之进行的结果。概率的出现只是由于实验者缺乏知识而已。值得注意的是，不能用这样的观点来解释量子力学中出现的令人困惑的，显然是非定域的概率！

为了展示这一点，让我们考虑一个和上面相像的情形，但是只有在A和B分隔得很开以后才决定对系统A测量的选择。似乎B的行为瞬息地受这个选择的影响！正是阿尔伯特·爱因斯坦、玻里斯·波多尔斯基和奈坦·罗森（1935）提出了这类似是而非的“EPR”型的“理想实验”。我将沿用大卫·玻姆（1951）提出的一个变种。从约翰·S.贝尔的一个杰出的定理（参阅*Bell 1987*，*Rae 1986*，*Squires 1986*）可以得到这样的推论，任何定域的“现实的”（例如隐变量，或“经典型的”）描述都不能给出正确的量子概率。

假定由一个在某一中心点自旋为零的粒子衰变产生两个半自旋的粒子——我将其称为电子和正电子（也即反电子），它们沿着相反方向做直线运动（图6.30）。由于角动量守恒，电子和正电子加起来的总自旋必须为零，这是因为原先中心粒子的角动量为零。这个实验的含义是，当我们在某一个方向测量电子的自旋，无论我们选择什么方向，正电子都在相反的方向上自旋！这两个粒子可以相隔几英里甚至1光年那么远。然而对一个粒子的测量的选择似乎瞬息地固定了另一个粒子的自旋轴。



图6.30　自旋为0的粒子衰变成两个自旋为1/2的粒子，一个电子E和一个正电子P。测量其中的

一个自旋为1/2的粒子的自旋，显然瞬息地决定了另一个粒子的自旋态

让我们看看量子的形式是如何地导致这一个结论的。我们用态矢量|Q〉来表达联合的双粒子的零角动量态，并发现下式成立：

|Q〉=|E↑〉|P↓〉-|E↓〉|P↑〉，

这里E是电子而P是正电子。这里的情形是按照自旋向上或向下的方向来描述的。我们发现，整个态应是自旋向上的电子和自旋向下的正电子以及自旋向下的电子和自旋向上的正电子的态的线性叠加。这样，如果我们在自旋向上或向下态的方向测量电子时，若发现电子自旋确实向上，则我们必须跃迁到态|E↑〉|P↓〉，这样正电子的自旋态必须向下。另一方面，如果我们发现电子自旋向下，则态跃迁到|E↓〉|P↑〉，这时正电子自旋向上。

假定我们现在选择其他的一对相反的方向，譬如向右的和向左的，而

|E→〉=|E↑〉|E↓〉，|P→〉=|P↑〉+|P↓〉；

|E←〉=|E↑〉-|E↓〉，|P←〉=|P↑〉-|P↓〉；

则我们发现（如果你愿意的话，可用代数检查一下！）

|E→〉|p←〉-|E←〉|p→〉=

（|E↑〉+|E↓〉）（|P↑-|P↓〉）-

（|E↑〉-|E↓〉）（|P↑+|P↓〉）=

|E↑〉|P↑〉+|E↓〉|P↑〉-|E↑〉|P↓〉-

|E↓〉|P↓〉-|E↑〉|P↑〉+|E↓〉|P↑〉-

|E↑〉|P↓〉+|E↓〉|P↓〉=

-2（|E↑〉|P↓〉-|E↓〉|P↑〉）=

-2|Q〉.

它（除了一个不重要的因子-2以外）和我们开始的态一致。这样，我们原先的态可同样合格地被认为是自旋向左的电子和自旋向右的正电子以及自旋向右的电子和自旋向左的正电子的态的线性叠加！如果我们要在向左或向右的方向上而不是向上或向下的方向上测量电子的自旋，这一个表达式就十分有用。如果我们发现电子的自旋向右，则态跃迁到|E→〉|P←〉，这样正电子的自旋就向左。另一方面，如果我们发现电子自旋向左，则态跃迁到|E←〉|P→〉。这样正电子自旋就向右。假定我们在任何其他方向上测量电子的自旋，其情景完全是相对应的：正电子的自旋态会立即跃迁到同一方向或者相反的方向上去，这要依赖于对电子测量的结果。

为何我们不能用一种类似的方法，以上述的从一个盒子中取出黑球和白球的例子，来作为我们电子和正电子的自旋的模型呢？让我们考虑一般的情形。我们现在不用黑球和白球，而用原先合在一起然后向两个相反方向运动的两台仪器E和P。假定不管E还是P都能对在任何方向进行的自旋测量作是或非的响应。对于选择任何的方向，其响应可以被仪器完全决定，或许仪器只产生概率的响应，其概率由该仪器所决定。但是，我们假定在分开之后，不管是E还是P都是完全相互独立的行为。

我们在每一边都有一台自旋测量仪，一台测量E的自旋，另一台测量P的自旋。假定在每台测量仪上都有自旋的三个方向的刻度，譬如E测量仪上的A、B、C和P测量仪上的A'、B'、C'。方向A'、B'、C'分别和A、B、C相平行。我们取A、B和C在平面上的相互夹角为120°（图6.31）。现在想象在每一边的不同的刻度将该实验重复多遍。有时E测量仪会记录上是（也就是自旋是在测量的方向A、B或C上），还有时候会记录非（自旋在相反方向）。类似地，P测量仪有时会记录是，有时会记录非。我们注意到实际量子概率必须具备两个性质：

（1）如果两边的刻度是同样的（亦即A和A'，等等），那么两个测量所产生的结果总是不同意（亦即，只要P测量仪记录非时，E测量仪就记录是，而且只要P给出是时E就为非）。

（2）如果将刻度盘随机地旋转并放置，两者完全相互独立，则两个测量仪同意或不同意的情况是等概率的。

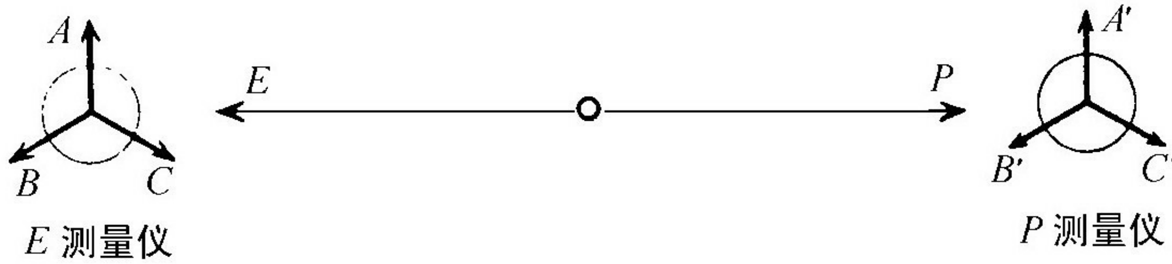


图6.31　EPR佯谬和贝尔定理的大卫·墨明简化形式，显示出在现实的定域的自然观点和量子理论的结果之间存在矛盾。E测量仪和P测量仪各自独立地具有测量它们各自粒子的自旋的三个方向刻度

我们容易看出，性质（1）和性质（2）是直接从我们早先的量子概率规则来的。我们可以假定E测量仪先动作。然后P测量仪发现粒子的自旋态，和E测量仪测量的结果相反。这样立即得到了性质（1）。为了得到性质（2），我们注意到，对于测量方向之间差120°的情形，如果E测量仪给出是，则P方向是和它所作用的自旋态夹角为60°；如果E给出非，则它和这自旋态夹角为120°。这样测量同意的概率为$\frac{3}{4}=\frac{1}{2}(1+\cos 60°)$，不同意的概率为$\frac{1}{4}=\frac{1}{2}(1+\cos 120°)$。所以，对于三个P刻度，如果E给出是，P也给出是的概率为$\frac{1}{3}=(0+\frac{3}{4}+\frac{3}{4})=\frac{1}{2}$，而P给出非的概率为$\frac{1}{3}(1+\frac{1}{4}+\frac{1}{4})=\frac{1}{2}$，亦即同意和不同意是等概率的。类似地，如果E给出非，情况也一样。这的确就是性质（2）。（参阅340页）

非常令人吃惊的是，性质（1）和性质（2）和任何定域的现实模型（亦即和所有能摹想到的这类仪器）都不协调！假定我们有这样的一个模型，E仪器必须准备好应付每一可能的A、B或C测量。我们注意到，如果只准备得到随机的答案，那么为了和性质（1）相符合，P仪器分别对于A'、B'和C'不能一定给出不同意的结果。的确，两台仪器必须对预先确定的准备好的三种可能的测量每种给出答案。例如，假定对于A、B、C这些答案分别为是、是、是；则右手的粒子就必须准备对于三个相应的右手刻度给非、非、非的答案。如果，左手准备的答案为是、是、非，则右手答案就必须为非、非、是。所有其他情况都在本质上和这些相似。现在让我们看看这是否和性质（2）相协调。做是、是、是/非、非、非的指定不是非常有助的，因为这时在所有可能的配对A/A'，A/B'，A/C'，B/A'等中有9种情形不同意，0种情形同意。关于其他情况，譬如是、是、非/非、非、是以及类似的情况又如何呢？有5种不同意，4种同意。（只要全部列举出来就能检验了：是/非、是/非、是/是、是/非、是/非、是/是、非/非、非/非、非/是，其中5种不同意，4种

同意。）这离开（2）的需要要近得多了，但还不够好，因为我们要求同意和不同意一样多！其他任何和性质（1）相协调的一对指定都会给出5比4（除了更坏的非、非、非/是、是、是情形，又给出9比0的答案）。不存在一组准备好的答案能产生量子力学的概率。因此，定域的现实模型必须被排除掉[14]！

光子实验：相对论的一个问题

我们应该问实际的实验是否支持量子力学的这些令人惊愕的预言。刚刚描述的精密的实验只是假想的，并没有被进行过。但是人们曾经利用一对光子的极化，而不是自旋为1/2的有质量的粒子的自旋进行过类似的实验。除了这个区别外，这些实验在本质上和上述的一样，除了有关的角度（由于光子的自旋为一，而不是一半）只是那些半自旋的粒子的一半。对光子的极化或偏振已在各种不同的方向组合上测量过，结果和量子力学的预言完全一致，而和任何定域的现实模型不协调！

迄今最精确和令人信服的实验结果是由阿铃·阿斯匹克特（1986）和他在巴黎的合作者得到的[15]。阿斯匹克特的实验还有另一个有趣的特点。以何种方法测量光子极化的“决定”是在光子完全飞走之后才做的。这样，如果我们认为存在从一个光子探测器跑到在相反一边的另一个光子探测器的非定域的，通知另外那个光子人们想要测量的偏振的方向的某种影响，则我们看到这种影响必须走得比光还快！任何和这事实相一致的量子世界的现实的描述，显然必须是非因果性的。这是在效应应该能比光传递得更快的意义上讲的。

但是，我们在上一章已经看到，只要相对论是正确的，用超光速发送信号就会导致荒谬（并和我们“自由意志”的感觉相矛盾等，参阅274页）。这肯定是对的。但是，在EPR类型实验中出现的非定域的“影响”，如果这样做的话就会导致荒谬，所以不能用以传递信息。（吉拉迪·雷米尼和韦伯在1980年详细地演示了这样的“影响”不能用于传递信号。）直到我们被告知实际是两种选择中的哪一种时，说一个光子“在垂直或水平”（或相反地说是在60°或者150°）方向偏振是没有用的。“信息”的这一部分（亦即不同的偏振方向）比光到达得更快（“瞬息”），而这两个方向中哪一个实际上被极化的知识，通过传递第一偏振测量的结果的通常信号，将更慢地到达。

在通常发送信息的意义上，虽然EPR类型的实验不和相对论的因果性发生冲突，它肯定和我们的“物理实在”的图像中的相对论精神相矛盾。让我们看看如何将态矢量的现实的观点应用到上述的EPR类型的实验（牵涉到光子）中去。当两个光子向外运动，态矢量描述作为单独单元的光子对的情形。没有一个光子单独地具有一个客观的态；量子态只适用于两个光子一起的情形。没有一个光子单独地有偏振方向；偏振是两个光子结合在一起的性质。当这两个光子中的一个偏振被测量时，态矢量就跃迁，使得未被测量的光子具有确定的偏振。当那个光子的偏振接着被测量时，将通常的量子规则应用到那个偏振态上去，就正确地得到了概率的值。用这种方式来看问题就得到了正确的答案；这正是我们通常应用量子力学的方法。但是，在本质上这是一种非相对论性的观点。因为这两个偏振的测量是称为类空分隔的。它表明任一测量都处于另一测量的光锥之外，正如图5.21中的点R和点Q的情形。两个测量哪个先发生的问题在实际上没有物理意义，它依赖于“观察者”的运动状态（图6.32）。如果观察者向右运动得足够快，则他认为右手的测量先发生；如果向左，则左手的测量先发生！但是，如果我们认为右手的光子先被测量，我们就得到了和认为左手光子先被测量的完全不同的物理实在的图像！（正是不同的测量引起了非定域的“跃迁”。）在我们物理实在的空间——时间图像——甚至是正确的非定域的量子力学的图像——和狭义相对论之间有本质上的冲突！这是一个严重的困惑，“量子的现实主义者”还不能予以解决（参阅*Aharonov and Albert 1981* ）。我在以后还要回到这问题上来。

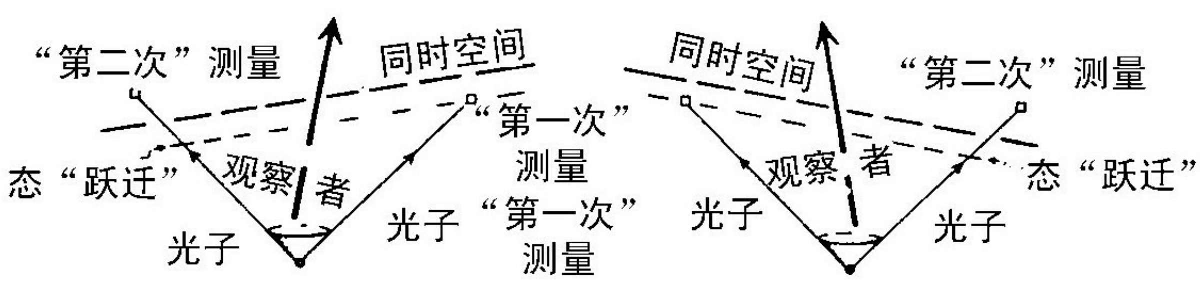


图6.32　在EPR实验中两个光子从一个自旋为零的态向相反的方向发射。两个不同的观察者形成“实在”的不一致的图像。向右运动的观察者判断态的左手部分在它被测量之前跃迁，这跃迁是由于右边的测量引起的。而向左运动的观察者的观点与此刚好相反

薛定谔方程；狄拉克方程

我在本章的前一部分提到了薛定谔方程。它是一个定义得很好的决定性的方程，在许多方面和经典物理的方程相当类似。该法则说，只要不对量子系统进行“测量”（或“观察”），薛定谔方程必须成立。读者或许会愿意看到它的实际形式：

$$i\hbar\frac{\partial}{\partial t}\ |\Psi\rangle = \mathrm{H}|\Psi\rangle。$$

我们会记得，$\hbar$ 是普朗克常数的狄拉克写法（h/2π），i=√−1，而作用到|Ψ〉上的算符∂/∂t（对时间的偏微分）就表示|Ψ〉对时间的变化率。薛定谔方程讲“H|Ψ〉”描述|Ψ〉是如何演化的。

但是“H”是什么呢？它是我们在前一章考虑过的哈密顿函数，但是这里有一个根本的不同！回顾一下经典哈密顿量是按照系统中的所有物理对象的各种位置坐标q_i 和动量坐标p_i 来表达的总能量。为了得到量子的哈密顿量，我们可取同样的表式，但是对每一处出现的动量p_i 要用微分算符“对q_i 的偏微分”的倍数取代。明确地讲，我们用$-i\hbar\partial/\partial q_i$ 来取代p_i 。我们的量子哈密顿量H就变成某种（经常是复杂的）牵涉到微分和乘法等的数学运算——而不仅仅是一个数！这有点像变魔术！但是它不仅仅是数学符咒，它是真正起作用的魔术！（应用这个过程从经典哈密顿量产生量子哈密顿量需要一点“艺术”，但是和其奇异的性质相比较，在这个过程中固有的、起作用的模糊之处是这么微小，真是令人印象深刻。）

薛定谔方程（不管H是什么样子的）是线性的，这是值得注意的重要之处。也就是说，如果|Ψ〉和|φ〉都满足该方程，则|Ψ〉+|φ〉或甚至任何组合ω|Ψ〉+z|φ〉都满足，这里ω和z为固定的复数。这样，薛定谔方程维持复线性叠加。两个可能的不同的态的（复）线性叠加不能仅仅由于U的作用而被“拆开”！这就是为何为了使只有一个选择存活下来，作为与U相分别的步骤R的作用是必需的。

薛定谔方程像经典物理中的哈密顿形式一样不是那么特殊的方程，而是量子力学方程的一般框架。一旦人们得到了合适的哈密顿量，态按照薛定谔方程演化的方式，使得|Ψ〉仿佛是服从于某种诸如麦克斯韦的

经典场方程的经典场。事实上，如果|Ψ〉描述一单独光子的态，那么薛定谔方程实际上成为麦克斯韦方程！单光子的方程刚好和整个电磁场的方程[54]完全相同。这一个事实是我们早先瞥见的单独光子的麦克斯韦场的类波动行为和偏振的缘由。另一个例子是，如果|Ψ〉描述单电子的态，则薛定谔方程就变成狄拉克著名的电子波动方程。这一个方程是他以伟大的创造性和洞察力于1928年发现的。

事实上，狄拉克电子方程必须和麦克斯韦方程以及爱因斯坦方程同列为物理学的伟大的场方程之一。为了使我们对之有深刻的印象，我就得必须引入令人眼花缭乱的数学观念。只要举一个例子就可以了，狄拉克方程中的|Ψ〉有一奇怪的“费米子”的性质，即在360°旋转下|Ψ〉变成-|Ψ〉，这一点我们早先已经考虑过了（336页）。狄拉克方程和麦克斯韦方程一道组成了最成功的量子场论——量子电动力学的基础。我们在下面简要地讨论它。

量子场论

所谓“量子场论”的学科是从狭义相对论和量子力学的观念的结合而产生的。它和标准（亦即非相对论性）的量子力学的差别在于，任何特殊种类的粒子的数目不必是常数。每一种粒子都有其反粒子（有时，诸如光子、反粒子和原先粒子是一样的）。一个有质量的粒子和它的反粒子可以湮没而形成能量，并且这样的对子可由能量产生出来。的确，甚至粒子数也不必是确定的；因为不同粒子数的态的线性叠加是允许的。最高级的量子场论是“量子电动力学”——基本上是电子和光子的理论。该理论的预言具有令人印象深刻的精确性（例如，上一章已提到的电子的磁矩的精确值，参阅199页）。然而，它是一个没有整理好的理论——不是一个完全协调的理论——因为它一开始给出了没有意义的“无限的”答案，必须用称为“重正化”的步骤才能把这些无限消除。并不是所有量子场论都可以用重正化来补救的，即使是可行的话，其计算也是非常困难的。

使用“路径积分”是量子场论的一个受欢迎的方法。它是不仅把不同粒子态（通常的波函数）而且把物理行为的整个空间——时间历史的量子线性叠加而形成的（参阅费恩曼1985年的通俗介绍）。但是，这个方法自身也有附加的无穷大，人们只有引进不同的“数学技巧”才能赋予意

义。尽管量子场论有毋庸置疑的威力和印象深刻的精确度（在那些理论能完全实现的很少情况），人们仍然觉得，必须有深刻的理解，才能相信它似乎是导向“任何物理实在的图像”[16]。

我应该澄清的是，由量子场论提供的量子理论和狭义相对论之间的一致性只是部分的——只对U过程——并且它具有相当数学形式的性质。量子场论甚至还未触及困难之处：对R过程中产生的“量子跃迁”（EPR类型实验留给我们的）做协调的相对论解释。此外，我们还没找到一个一致的或可信的引力量子场论。我将在第8章提议，这些问题也许不是完全相互无关的。

薛定谔猫

最后让我们回到从一开始描述就尾随我们的问题。我们为何从未见到经典尺度现象的量子线性叠加，诸如板球同时处于两个地方？究竟是什么东西使得构造测量仪器的原子的某种形态能用过程R来取代U？任何测量仪器自身无疑是物理世界的一部分，它是由那些量子力学的构件制配而成，它的行为是被设计来作此探索的。为何不将测量仪器和被考察的物理系统一起作为合并的量子系统来处理，如果这样就不牵涉到神秘“外界”的测量。这合并的系统应简单地按照U来演化。但是，果真如此吗？U在合并系统的作用是完全决定性的，并没有R类型的概率不确定性卷入到合并系统并对自身进行“测量”或“观察”的余地！这里存在一个明显的矛盾，在厄文·薛定谔（1935）引入著名的理想实验——薛定谔猫的矛盾中变得特别写实。

想象一个封闭的容器，它制造得如此完美以至于没有任何向内或向外的影响能通过容器壁。想象在容器里有一只猫，并且还有一台能被某量子事件触发的仪器。如果该事件发生，该仪器打碎装着氰化物的药瓶，并将猫毒死。如果该事件没发生，则猫继续活着。在薛定谔原先的设计中，量子事件为放射性原子的衰变。让我稍作修正，并把光子触发光电管作为我们的量子事件。在这里光子是由某个处于预先确定状态的光源发出，然后由半镀银的镜子反射下来（图6.33）。镜面的反射将光子波函数分裂成两个分开的部分，由该镜子使之一部分反射而另一部分穿透。光子波函数的被反射部分聚集在光电管上，这样如果光子被光电管所记录，它就是被反射的。这种情形下，氰化物就流出来，猫就被毒

死。如果光电管没有记录，光子就穿透过半镀银的镜子而到达后面的墙上，猫就存活。

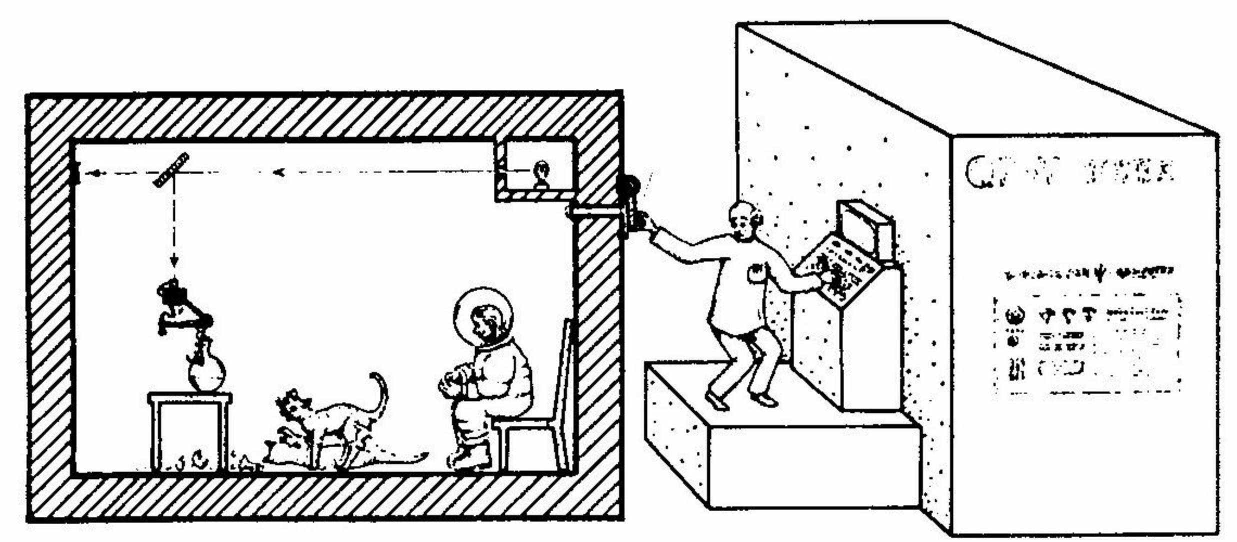


图6.33　薛定谔猫——以及附加物

从处在容器内的（有点危险的）一个观察者的观点，这的确是在那里所发生的描述。（我们最好为此观察者提供合适的防护服！）或者光子被反射，因为光电管“观察到”并记录到，猫被毒死；或者光子穿透过，由于光电管没有“观察到”并没有记录，猫是活的。实际上，两者必居其一：R起了作用，每一种可能性的概率为百分之五十（因为它是一面半镀银的镜子）。现在，让我采用处于容器之外的物理学家的观点。我们可以认为，在容器被封之前他已知内部的初始态矢量。（我不是指在实际上他能知道，而是量子理论没有说在原则上不能让他知道。）根据外面的观察者，在实际上没有进行“测量”，这样整个态矢量必须按照U进行。光子由处于预定的状态的源中发出——两个观察者在这一点上是一致的——它的波函数分成两束，譬如讲每一部分光子的幅度均为$1/\sqrt{2}$（这样平方模就给出1/2的概率）。由于这整个系统被外界的观察者当作单独的量子系统来处理，不同选择之间的线性叠加必须一直保持到猫的尺度。光电管记录到和没有记录到光子的幅度各为$1/\sqrt{2}$。在这种态下两种选择都必须存在，在量子线性叠加中权重相同。根据外面的观察者，猫是处于死和处于活的线性叠加态！

我们真的会相信这种事吗？薛定谔本人清楚地表示他不相信。他论证道：量子力学的U规则实际上不能适用于像猫这么大、这么复杂的东西上。在这过程中薛定谔方程一定出了什么差错。当然薛定谔有权利用

这种方式来评论他的方程，但是我们并没有分享到这种特权！相反地，大量（也许大多数）物理学家宁愿坚持，现在有如此大量的实验证据支持U—没有一个人反对之——甚至在猫的尺度下，我们没有什么权利去抛弃这类演化。如果这一点被接受，我们就似乎被导致到物理实在的非常主观的观点。对于这外面的观察者，猫的确是处于活和死的线性组合中，只有当容器最后被打开后猫的态矢量才坍缩成其中的一种选择。另一方面，对于在里面的（适当防护的）观察者，猫的态矢量坍缩得早得多，而外面观察者的线性叠加和他不相干。

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\{|死\rangle + |活\rangle\}$$

态矢量似乎“完全处于精神之中”！

但是，我们真能采用态矢量的这种主观观点吗？假定外面的观察者做了某些复杂得多的事，而不仅仅是“窥视”该容器。假定他首先从他得到的容器内部的初始态的知识，使用他能得到的一台大型计算设备，由薛定谔方程计算3出容器内的态应实际上是什么样的，得到了（正确的！）答案|Ψ〉。这里|Ψ〉的确是上述的死猫和活猫的线性叠加。然后他进行一个特殊的实验，把这个态|Ψ〉和所有与之正交的态鉴别开来。（根据前述的量子力学规则，他在原则上可以进行这样的实验，尽管在具体实现时会遭遇到极大的困难。）“是的，它是处于态|Ψ〉”和“不，它处于与|Ψ〉正交的态”的两种结果的概率分别为百分之百和百分之零。特别是，态|〉=|死〉-|活〉的概率为零，它是和|Ψ〉正交。|〉作为实验结果的不可能性只能是因为两个选择|死〉和|活〉共存并相互干涉而引起的。

如果我们稍稍调整光子的路径长度（或镀银的量），使所得到的态不是|死〉+|活〉，而是别的组合，譬如|死〉-i|活〉，等等。所有这些不同的组合在原则上都具有不同的实验后果！所以它甚至“不仅”是某种会影响我们的可怜的猫的死亡和存活的共存的事体。所有不同的复组合都是允许的，它们在原则上应能互相被区分开来！然而，对于容器内的观察者，似乎所有这些组合都是无关紧要的。猫或者是活的，或者是死

的。我们如何理解这种偏离呢？我将简要地指出一些关于这些（以及相关的）问题的不同观点，—虽然毫无疑问地，我将不会完全公平地对待它们！

现存量子理论的不同看法

首先，在实现诸如将态|Ψ〉与任何和|Ψ〉正交的态区分开来的实验中存在着明显的困难。毫无疑问地，在实际上，这种实验对于外面的观察者而言是不可能的。特别是，甚至在他计算|Ψ〉将来实际上应是什么样子之前，他需要知道（包括内部观察者的）整个内容的态矢量！然而，我们要求这个实验不仅在实际上，而且在原则上不可能实现，否则我们就没有权利从物理实在中移走态|活〉或态|死〉中的一个。麻烦在于，量子理论的现状并没有在"可能的"测量和"不可能的"测量之间划上一道清楚界限的法规。也许应该存在这样清楚的区别。但是，理论的现状不允许这种东西。引进这种区别就会使量子理论改观。

其次，一种相当普遍的观点认为，如果我们充分地考虑环境的影响，则困难就会被消除。的确，要使系统完全和外界隔离在实际上是不可能的。只要外界的环境牵涉到容器内的态，则外部观察者就不能认为系统是由一个单独的态矢量来描述。甚至他自己的态和这系统以一种复杂的方式相关联。况且，还有大量的不同粒子纠缠以及一直弥散到宇宙中越来越远的，包括极大量自由度的不同可能的线性组合的效应。不存在一种可行的方式（譬如靠观察适当的干涉效应）把这些复线性组合从仅仅为概率加权的选择中区别出来。这甚至不必是把系统和外界隔离开来的问题。猫本身牵涉到巨大数量的粒子。这样死猫和活猫的复线性组合可以像简单的概率混合那样处理。然而，我本人认为这根本不是令人满意的。正如对付前面的观点一样，我们可以问在哪一阶段可以正式认为"不可能"得到干涉效应——使得可以宣布说复线性叠加的幅度平方模提供了衡量"死"和"活"的概率？甚至如果世界的"实在"在某种意义上"在实际上"变成一个实数概率权重，如何将它只分解成这种或那种选择？在仅仅依赖演化U的基础上，我看不到实在如何将两种选择的一个复（或实）线性叠加变换成其中的这样一种选择。我们似乎被逼回到世界的主观观点上去！

有时人们采取这样的观点，复杂的系统实际上不应该由"态"而应由

所谓的密度矩阵的推广来描述（*von Neumann 1955* ）。这些同时牵涉到经典概率和量子幅度。事实上，许多不同的量子态被一起用来代表实在。密度矩阵是有用的，但是它们自身不能解决量子测量深刻而可疑的症结。

人们也许同意，实际的演化是决定性的U，但在了解该组合系统的量子态究竟是什么时牵涉到的不确定性引起了概率。这可认为是关于概率起源的非常“经典的”观点——它们全部是从初始态的不确定性引起的。人们可以想象，微小的初始态的差别会产生演化中的巨大差别。正如经典系统会产生“混沌”一样（譬如，天气预报，参阅第5章224页）。然而，单由U本身不会产生这种“混沌”，因为它是线性的：在U的作用下，人们不想要的线性叠加被一直维持着。要把这种叠加归结成这种或那种选择，U本身做不到，需要某种非线性的东西。

作为另一种观点，我们也许注意到了这个事实，在薛定谔猫的实验中唯一和观察结果完全明确的偏差似乎是由于有意识的观察者引起的，一个（或两个）在容器里面和另一个在外面。也许复量子叠加定律不能应用于意识！欧根·P.维格纳（1961）为此观点提出了一个粗糙的数学模型。他提议，薛定谔方程的线性也许对于有意识的（或仅仅是“活”的）本体无效，它由某种非线性的步骤所取代，由此被归结成两种选择中的一个。读者或许会认为，由于我在寻求某种量子现象在我们意识思维中的作用——我们的确如此，我倒最为同情这种可能性。然而，我一点也不喜欢它。它似乎会导致世界实在的非常不均衡的使人烦恼的观点。宇宙中意识栖息存在的角落可以说是非常稀少并相隔得非常远，依此观点，复线性叠加只在那些角落归结成实际的选择。情况也许是这样，对我们来说，其他这样的角落和宇宙的其余部分显得相同，因为不管我们自身看到（或观察到）什么，由于我们意识的行为使它“归结成选择”，而不管是否之前已经归结成这个样子。若果真如此，这种巨大的失衡会给世界的实在性提供一个非常使人烦忧的图像，而要我作为其中一员只能非常犹豫地去接受它！

还有一种相关的称作参与宇宙的观点（由约翰·A.惠勒在1983年提出），将意识的作用推向一个（不同的）极端。例如，我们注意到，这一个行星上的意识生命的演化是由于不同时期的适当的沧桑巨变。这些被设想为量子事件，所以它们只在线性叠加的形式中存在，直到它们最后导致意识生命的演化——其存在完全依赖于正确的巨变“在实际上”发

生！依此观点，正是我们自身的存在把我们的过去变戏法为存在。此图像中的逻辑循环的矛盾引起人们的一些注意，但我自己感到这种观点困难重重，并且几乎是不可信的。

另外一种本身是逻辑性的，但是提供出同等奇怪图像的称为多世界的观点。这是休斯·埃维勒特三世首次公开提出的（1957）。按照多世界解释，R根本从未发生过。实在的态矢量的全部演化被认为总是由决定性的过程U所制约的。这意味着可怜的薛定谔猫和容器中的受防护的观察者的确应该存在于一种复线性组合之中，猫处于某种活和死的叠加态中。然而，死的状态是和内部观察者意识的一种态相关，而活的与另一状态相关（并且假定，部分地和猫的意识相关——并且当这些内容呈现给外界观察者时，最终也和他相关）。每一观察者的意识被看作“分裂”，这样现在他存在两次，每一次他的情形都有不同的经验（也就是，一次看到死猫，另一次看到活猫）。的确不仅是一个观察者，他所居住的整个宇宙都在他对宇宙所进行的每一“观察”中分裂成两个（或更多个）。这种分裂不断地发生——不仅仅是由于观察者进行的“观察”，而且还一般地由于量子事件的宏观的放大——这样使得这些宇宙“分枝”疯狂地蔓延。的确，每一种不同的可能性都会在某种巨大的叠加中共存。这肯定不是最经济的观点，但是我本人反对它的原因并不是这种不经济。特别是，我看不出为何意识只能知晓线性叠加的“一”个选择。是有关于意识的什么东西使人们无法“知晓”令人焦虑的死猫和活猫的线性叠加呢？我似乎觉得在多世界观点和人们实际观察到的之间相符合之前必须先有关于意识的理论。在宇宙的“真正”（客观）态矢量和我们要实际“观察”到的之间我看不到什么关系。有人断言，R的“幻像”在某种意义上能在这图像中被等效地导出，但我认为这一断言不成立。要使这种方案可行，人们至少需要进一步的要素。依我看来，多世界观点并没有在实际上触动量子测量的真正的困惑，而自身却引进了许多问题。（比较*Witt and Graham 1973* 的讨论。）

何处出了差错

就量子力学的理论现状而言，任何解释上的困惑总是以这种或那种面目出现而挥之不尽。让我们简略地复习一下标准的量子理论在实际上告诉我们应如何描述世界，尤其是和这些令人困惑的问题之间的关系。然后我们向自己提出这样的问题：我们将往何处去？

首先，我们知道只能把量子理论的描述有意义（有用）地应用到分子、原子或亚原子粒子的所谓量子水平上去。但是，只要在不同的可能性之间的能量差保持非常小时，也能在大尺度下应用。在量子水平上，我们应该把这种“选择”当作可共存的东西来处理，以一种复数权重来叠加。我们用以加权的复数称为概率幅。每一不同的复加权选择的总体定义一个不同的量子态，而任一个量子系统必须用这样的量子态来描述。以自旋的情况作例子最为清楚了。对于什么是构成量子态的“实际的”选择以及什么仅仅是选择的“组合”，我们无可奉告。无论如何，只要系统仍处于量子水平，量子态就以完全决定性的形式演化。由重要的薛定谔方程制约的过程U即是这种决定性的演化。

当不同量子选择的效应被放大到经典水平，使得选择之间的差别足够大到我们可以直接感知，那这样的复权重叠加似乎不再维持。相反地，复幅度的平方模被形成（也即把它们在复平面上的位置离开原点的距离取平方），而现在这实数扮演问题中选择的实际概率的新角色。只有其中的一个选择依照过程R（称为态矢量的减缩或波函数的坍缩，完全和U不同）在物理经验的实在中存活。量子理论的非决定性正是在这里也仅是在这里被引进来。

人们也许可以有力地为量子态提供一个客观的图像辩护，但是它是复杂的，甚至有些使人觉得似是而非。当有若干个粒子参与时，量子态（通常）会变得非常复杂。单独粒子自身不再有它们自己的“态”，而是处于和其他粒子相缠结的复杂的相关状态中。当在一个区域“观察”一个粒子时，也就是它触发了某种效应使之放大到经典水平，那么必须祈求R—但是这显然同时地影响其他和该粒子相关的所有粒子。爱因斯坦、波多尔斯基和罗森（EPR）类型的实验（譬如在阿斯匹克斯实验中，由一个量子的源向相反方向发射出一对光子，然后在相隔几米的距离下分别测量它们的偏振）对这些量子物理困惑的，却又是根本的事实给出了清楚的观察结果：它是非定域的（使得阿斯匹克斯实验中的光子不能被当成分开的独立的本体来处理）！如果R被认为是一种客观方式的作用（它似乎为量子态的客观性所隐含），那就相应地违背了狭义相对论的精神。看来不存在能和相对论要求相一致的（正在减缩的）态矢量的真正客观的空间——时间描述。然而，量子理论的观察效应不违反相对论。

量子理论在关于何时和为何R实际上（或显得？）发生的问题上保

持缄默。并且，它本身并没有适当解释为何经典水平的世界“显得”经典。要知道“大多数”量子态根本不像经典态！

何处出了差错？我相信，人们必须认真地考虑量子力学在应用于宏观物体时根本错了的可能性，或者定律U和R只不过是提供极为近似某种更完全的，但还未发现的理论。正是这两个定律结合在一起提供了现在理论而不光是U所享有的与观察的美妙的符合。如果把U的线性推广到宏观世界去，我们就必须接受板球等不同位置（或不同自旋等）的复线性叠加的物理实在。常识告诉我们，这不是世界真正行为的方式！经典物理的描述的确为板球提供了很好的近似。它们具有定义得相当好的位置，并没有出现量子力学线性定律所允许的同时处于两处的情况。如果过程U和R为更广泛的定律所取代，则新定律不像薛定谔方程那样，它具有非线性的特征（因为R自身非线性地起作用）。有些人持反对态度，他们完全正确地指出，标准量子理论深奥优美的数学性是来自于它的线性。但是我感到，如果量子理论在将来不遭受到一些根本的改变，那是不可思议的——对于某种东西它会变成线性只能是一种近似。牛顿的优雅而有力的万有引力理论要大大地归功于这一个事实，理论中的力以线性的方式相加。然而，和爱因斯坦广义相对论相比，这种线性只是（虽然是极好的）近似——爱因斯坦理论的精巧甚至超过了牛顿理论！

我毫不犹豫地相信，量子理论矛盾的解决在于我们找到一个改善的理论。虽然这也许不是传统的观点，但也不是毫无传统可言。[许多量子理论的创始者也有这种想法。我是指爱因斯坦的观点。薛定谔（1935）、德布罗意（1956）和狄拉克（1939）也认为此理论是临时的。]但是，甚至如果人们相信此理论是要进行某种修正，而应该如何进行修正的方式还要受到巨大的限制。也许某种“隐变量”观点最终会变成可接受的。但是，由EPR类型的实验展示的非定域性对任何在通常空间——时间中能安然发生的世界“现实的”描写都构成了严重的挑战——这正是依照相对论原始所提供给我们的特殊类型的空间——时间——所以我相信需要更多得多的激变。况且，从未发现量子理论和实验之间的任何种类的偏离——当然除了人们把板球线性叠加态的不存在当成反例之外。依我自己的观点看，不存在线性叠加的板球正是相反的证据！但是这对它本身并没有什么大帮助。我们知道，量子定律支配着亚微观水平的东西，而经典物理支配着板球水平的东西。为了看到量子世界如何和经典世界合拢，在它们中间的某个地方我们必须对新的定律有所理解。我还相信，如果想理解思维的话我们必须理解这种新的定律。我相

信，为了所有这一切，我们必须寻求新的线索。

在本章的量子理论描述中，我完全采用传统的办法，虽然也许比通常更加强调几何和“现实性”。我将在下一章寻找某些必需的线索——我相信它能为改善量子理论提供某些暗示。我们从家乡开始旅行，但将被迫浪迹天涯。我们必须探索空间的极遥远处，并且要回溯到时间最初的起点！

第7章　宇宙论和时间箭头

时间的流逝

体验时间进展的感觉是我们知觉的中心。我们似乎从确定的过去向未定的将来不断前进。我们觉得过去的已经完结了，它是不可改变的，它在某种意义上还在“那里”。我们现在关于它的知识来自于我们的记录、我们记忆的痕迹以及从这些推导而来的东西。但是，我们从未怀疑过去的“实在性”。过去的那个样子也只能是这样了。发生过的事情已经发生过了，不管是我们还是任何人做任何事情都无法改变它！另一方面，将来似乎还是未定的。它可以这样也可以那样。或许这种“选择”完全是由物理定律所决定，或许一部分由我们自己（或上帝）所决定；但是似乎这种“选择”仍然有待于进行。它似乎仅仅是任何未来的“实在”都可以在实际上归结于它的潜势力。当我们有意识地感觉到时间的流逝时，广漠而表面上不确定的将来的最急切部分连续地变成为现实，并因此进入僵死的过去。有时我们会感到，我们甚至对特殊潜在的未来选择的某种影响独自“负责”，这种选择事实上已被实现，并成为过去的永恒实在。我们更经常觉得，当确定的过去疆域无情地吞噬未定的将来时，自身只是一个无助的旁观者——也许还要庆幸自己对这一切不必负责任。

但是，正如我们所知道的，物理告诉我们的却是另一回事。所有成功的物理方程都在时间上是对称的。它们在时间的任何方向上使用都显得一样。在物理学上，将来和过去似乎是平权的。牛顿定律、哈密顿方程、麦克斯韦方程、爱因斯坦广义相对论、狄拉克方程、薛定谔方程——如果我们颠倒时间方向（用-t来取代代表时间的坐标t），所有这些方程在实质上都不变。全部经典力学以及量子力学的U部分都是完全时间可逆的。现在存在一个问题，量子力学的R部分在实际上是否时间可逆的。这个问题将是下一章论证的中心。此刻，让我们首先避开这个问题，并把它当作这个课题的“传统智慧”，也就是不管其初看起来怎样，R的动做也应该被认为是时间对称的（参阅*Aharonov, Bergmann, and Lebowitz* 1964）。如果我们接受这些，似乎就必须环视四周，看看是否在他处能找到物理定律断言的过去和将来的差别之所在。

我们研究这个问题之前，必须考虑在我们的时间感觉和现代物理理论教导我们所相信的东西之间另一个令人困惑的偏离。根据相对论，根本就没有什么叫作“现在”的东西。我们所能得到和这最接近的概念是（正如258页的图5.21）观察者在空间——时间中的同时空间，但是它依赖于观察者的运动！一个观察者的“现在”和另一观察者的不同[1]。关于空间——时间中的两个事件A和B，第一位观察者U会认为B属于固定的过去，而A属于未定的将来；而对于第二观察者V可变为A属于固定的过去，而B属于未定的将来！（图7.1）。只要A和B中的任何一个事件是确定的，我们就不能完全有意义地断言另一个事件是否仍是未定的。

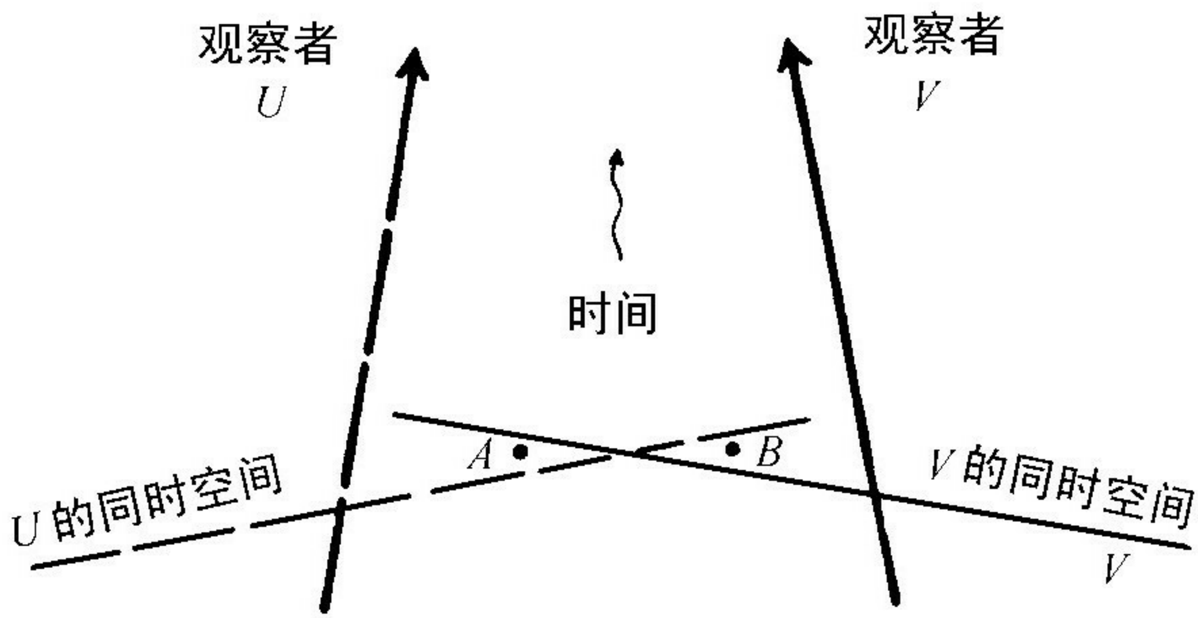


图7.1　时间真能流逝吗？从观察者U看来，B在“固定的”过去，而A还处于“未定的”将来，观察者V的观点刚好相反

回想一下259—260页的讨论以及图5.22。两人在路上相遇，按照其中一人，仙女座大星云太空舰队已经启程，而另一人却认为，还没有决定是否实际进行这次航行。那个已经决定的结果怎么还会有某种不确定呢？如果对于其中一个人而言决定已做出，那很清楚不能再有任何非确定性。太空舰队的启程已是不可避免。事实上他们中没有任何一个人知道太空舰队的发射。他们将来只能在地球上的望远镜观测揭示了舰队的确已在航程中时才知道。然后，他们可以回到原先邂逅之处[2]，并且得出结论，在那个时刻，按照其中一人，这个决定于未定的将来才做，而对于另一人，决定已在固定的过去做过。那时关于未来是否确有任何

未定之处？或者是否两人的未来都已被“固定了”？

情况似乎变成，如果任何事情完全确定，则整个空间——时间应该的的确确是确定的！不可能有“未确定的”未来。整个空间——时间必须是固定的，没有任何不确定的疆域。的确，这似乎正是爱因斯坦自己的结论（参阅Pais　1982，P444）。此外，根本就没有时间流逝。我们只有“空间——时间”——并且根本就没有正在被确定的过去无情侵占的未来疆域！（读者也许会诧异量子力学的“不确定性”在所有这些中扮演什么角色。我将在下一章回到量子力学引起的这一问题。此刻，最好只按照纯粹经典的图像来思考这一切。）

依我看来，在我们关于时间流逝的意识感觉和我们关于物理世界的实在的（超等精密的）理论所作的断言之间存在着严重的偏离。假定（正如我所相信的）知觉的更基础的某种东西一定能在和某种物理的关系中得到理解的话，则这些偏离必须在实际上告诉我们这种物理的一些深刻的内容。看来不管什么物理在起作用，它至少必须有一根本的时间反对称要素，也就是说它应该能把过去和将来区分开来。

如果物理的定律不能区分将来和过去——并且甚至连“现在”这个概念和相对论都不能和谐相处——那么究竟何处可以寻找到和我们自以为理解世界的方式更一致的物理定律呢？事实上，事情并非像我似乎要表明的那样具有这样大的偏离。我们的物理理解除了仅仅是时间演化的方程以外，还包含有牵涉到时间不对称的重要部分。其中最重要的是热力学第二定律。我们先要对这一个定律有所了解。

熵的无情增加

想象把一杯水放在桌子的边缘上。稍微推一下就会落到地面上去——无疑地会被打碎成许多碎片，水会溅到相当大的面积上，或许会被地毯吸收，还会流到地板的缝隙中去。我们这一杯水在这里只不过忠实地遵循着物理的方程罢了。牛顿的描述即已足够。杯子和水中的原子独立地遵守牛顿定律（图7.2）。现在让我们把这图像在时间的相反方向表演。由于这些定律的时间可逆性，这些水可以一样容易地从地毯和地板缝隙中流出，流进一个从许多碎片拼凑而成的玻璃杯中，这整体从地板上刚好跳跃到桌子的高度，然后停在它的边缘上。正如杯子落下打碎

的过程一样，所有这一切又都和牛顿定律相符合。

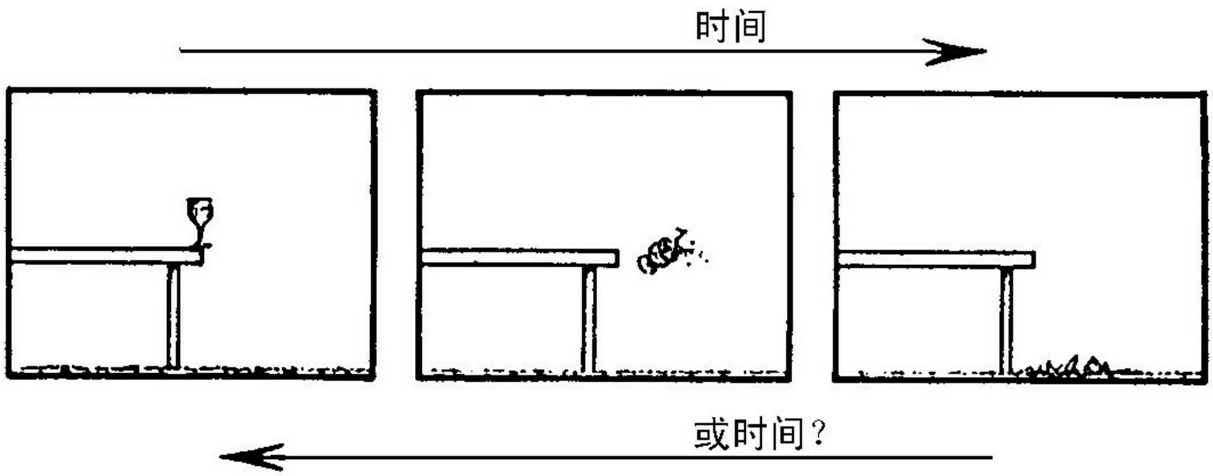


图7.2　力学定律是时间对称的；但是由右图到左图这样景象的时间顺序从未实现过，而由左图到右图则是司空见惯的

读者也许会问使杯子从地板上升到桌子上去的能量从何而来。那没有问题。不可能有能量的问题，因为在杯子从桌子落下时，从下落得到的能量必须跑到某处去。下落杯子的能量事实上变成热。在杯子摔到地面的时刻，杯子碎片、水、地毯和地板的原子会以一种比以前更快一些的杂乱的方式运动。也就是说，玻璃片、水、地毯和地板会比这发生之前仅仅变得稍热一些（不管蒸发引起的可能的热丧失——但是在原则上，那也是可逆的）。由于能量守恒，这热刚好等于这杯水从桌子上落下时的能量损失。所以，这些热能也刚好是足以使玻璃杯重新举到桌子上的能量！注意，在我们考虑能量守恒时把热能也计入是很重要的。把热能也包括进去的能量守恒定律称为热力学第一定律。由牛顿力学推导而来的热力学第一定律是时间对称的。第一定律并不以任何方式限制玻璃和水，从而排除碎片聚集成杯子，并且充满水后奇迹般地跳回到桌面上的可能性。

我们从未看到这类事情发生的原因是，在玻璃碎片、水、地板和地毯中的原子的“热”运动全是极其紊乱的，所以大部分原子都在错误的方向上运动。为了聚集玻璃碎片并收回所有溅开的水，而且最后优美地跳回到桌子上，必须以不可思议的精确度把它们的运动协调起来。可以肯定的是，这样协同的运动实际上是不存在的！只有极其侥幸地，也就是如果真有这样的“魔术”发生的话，才会有这种协同。

然而沿着时间的另一方向，这种协同运动则是司空见惯的。假定在物理状态的某种大尺度变化发生（这里是玻璃杯被打碎，水流走）之后而不是之前，粒子以协同的方式运动，我们并不把这些认为是侥幸。在此事件以后，粒子的运动的存在必须是高度协同的；由于这些运动具有这类性质，所以如果我们以完全精确的方式去颠倒每一个别原子的运动，则结果正是集中碎片，充满水并把水杯刚好举到出发之处所需要的行为。

把高度协调一致的运动看作大尺度变化的效应而不看作原因的观点是可以接受的并且是熟悉的。然而“原因”和“效应”两词需要面对时间反对称的问题。通俗地讲。我们已习惯于在原因必须先于效应的意义上应用这些术语。但是要想理解在过去和将来之间物理上的不同，就必须非常警惕不让我们日常的关于过去和将来的感觉无意识地注入讨论中去。我必须警告读者，要避免这样是极其困难的，但是我们必须强制自己这样做。我们必须尽力地这样使用词句，即在过去和将来的物理差异上不偏不倚。相应地，如果情势被认为刚好是合适的，我们就必定允许自己把事件的原因放在将来，而把效应放到过去！经典物理的决定性的方程（或量子物理的U过程）对于未来方向的演化并没有什么特权。它们可以一样好地适用于向过去方向的演化。未来之决定过去犹如过去之决定未来。我们可以用某种任意的方式指明系统在将来的某一个状态，并用之来计算过去应该是什么样子的。如果我们允许在时间的正常未来方向演化方程时，把过去当作“原因”，而把将来当作“效应”；则在时间的过去方向上，我们就可以应用演化方程的同等有效的步骤，并且显然地应该把将来当作“原因”，而把过去当作“效应”。

然而，在我们使用“原因”和“效应”的术语时牵涉到其他的某些东西，这根本就不是哪个事件发生在过去、哪个发生在将来的问题。让我们想象一个假想的宇宙，而且我们自己宇宙中的时间对称的同样的经典方程可适用于它。但是，在这宇宙中人们熟悉的行为（例如，一个玻璃杯被打碎，水流走）和这些行为的时间反演的发生共存。随同我们比较熟悉的经验，假定有时玻璃碎片真的聚集起来，神秘地充满了流走的水，然后又跳回到桌上去；还假定，有时搅拌煮熟的鸡蛋魔术般地恢复回来并最后飞回到打碎的蛋壳里，蛋壳完好地聚集起来，并把它新得到的内容封好；从溶解在甜咖啡中的糖会形成一块方糖，并自动地从杯子里跳回到某人手中。如果我们生活在这类事为司空见惯的世界中，我们肯定不会把这类事件的“原因”归结成奇异的有关单独原子的相关行为的

不可能的机遇，而是认为是某种“目的论效应”。由于这种效应，自装配的物体有时力求得到所需要的某种宏观的结构。“看！”我们会说，“它正在重新发生。那团乱七八糟的东西正把自己聚集成另一杯水！”我们会毫无疑问地认为，原子的目标是如此之精确，因为这是产生桌子上的一杯水的方式。桌子上的杯子变成“原因”而地面上显得杂乱的一团原子是“效应”——尽管这个“效应”在时间上比“原因”发生得更早。类似地，在搅拌煮熟的鸡蛋中的原子的精细组织的运动不是向聚集的鸡蛋壳跳回的“原因”，而是未来所发生的“效应”；糖块不是“因为”原子以非凡的精度运动，而是由于某个人——显然是在将来——要把糖块抓到手里，所以才集合起来并从杯子里跳出来！

当然，在我们的世界中看不到这类事的发生——或者可以更好地表达成，我们没看到这些事和那些正常类型的事共存。如果所有我们看到的都和上述的那样反常，则我们不会有任何问题。只要在我们所有的描述中把“过去”和“将来”，“以前”和“以后”等术语互相交换一下就可以了。可以认为时间沿着和原先认定的相反的方向前进，那个世界就可描述成和我们自己的世界一样。然而，我在这里摹想另一种不同的可能性——水杯的破碎和聚集能共存。在这样的世界中，我们不能仅仅靠改变时间进展的方向的习惯方法来恢复我们所熟悉的描述。当然，我们的世界刚好不是那样子，为何不是那样子？为了着手理解此事实，我要求你尝试想象这样的一个世界，并惊异我们会如何描述其中发生的事情。我要求你接受，在这样的一个世界中，我们一定能把粗糙的宏观的东西——诸如一满杯水，没有碎的蛋，手中的方糖——描述成提供的“原因”，而将详细的，或有精密关联的个别原子运动当作“效应”，而不管“原因”是否处于效应的将来或过去。

为何在刚好我们生活其中的世界中，在实际上原因总是超前于效应；或换一种讲法，为何准确协同的粒子运动总是在某种物态的大尺度变化之后而不是之前呢？为了对这类事物有更好的描述，我必须引进熵的概念。粗略地讲，系统的熵是其呈现的无序的量度。（以后我会表达得更精确一些。）这样，碎玻璃杯和地板上溅开的水，是比桌子上完好的一满杯水具有更高的熵的态；搅拌的鸡蛋比新鲜的未打碎的蛋具有更高的熵；甜咖啡比淡咖啡以及未溶解的糖块的熵更大。低熵态似乎是某种以明显的方式“被特别地安排好”，而高熵态却没有那么“被特别地安排”。

当我们谈到低熵态的“特殊性”时，很重要的一点是要意识到，我们指的是显明的特殊性。因为，在一个更微妙的意义上，这些情形下的高熵态，由于个别粒子运动的非常精密的协调，正和低熵态一样地是“被特别地安排的”。例如，在打碎杯子后流到地板缝隙中的水分子的似乎随机的运动其实是非常特殊的：其运动是如此之精密，如果它们所有都刚好颠倒过来，则原先的低熵态也就是桌子上的完好的、装满水的杯子就会被恢复。（情况必定如此，由于所有这些运动的反演刚好简单地对应于时间方向的反转——依此杯子会聚集好，并跳回到桌子上去。）但是，所有水分子的这种协调的运动并非我们称为低熵的那种“特殊性”。熵是指显明的无序性。存在于粒子运动的精确的协同的有序不是显明的有序，故不能用以降低系统的熵。所以，流出的水中的分子的有序性在这种方式中不能算数，它的熵是高的。然而，在完好的一杯水的显明的有序给出了低的熵值。这里表明的是这样的一个事实，即粒子运动只有少数几个可能的形态和一个完好装满水的杯子的显明形态相一致；相对来说，有更多得多的运动与地板缝中稍微加热的流水的显明形态相一致。

热力学第二定律断言孤立系统的熵随时间增加（或对于一个可逆的系统保持常数）。我们不能把协同的粒子的运动当作低熵。如果算的话，根据此定义，系统的“熵”就会永远是常数。熵概念只能指的确是显明的无序性。对于一个和宇宙的其余部分隔离开的系统，它的总熵增加。所以，如果它从某种显明的组织好的状态出发的话，该组织在过程中就会被腐蚀，而这些显明的特征就转化成“无用的”协同的粒子运动。第二定律似乎是一个绝望的裁决，因为它断言存在一个无情和普遍的物理原则，它告诉我们组织总是被不断地损坏。我们将来会看到，这个悲观的结论并非完全合适！

什么是熵

但是精确地讲，物理系统的熵应是什么呢？我们看到了它是显明无序的某种测度。但是，由于我这样不精密地使用诸如“显明”和“无序”的字眼，熵的概念实在还算不上一个清晰的物理量。第二定律还有另一方面似乎表明熵概念中的不精确的因素：只有所谓的不可逆的系统熵才实际上增加，而不仅仅是保持常数。“不可逆”是什么含义呢？如果计入所有粒子的细节运动，则所有系统都是可逆的！我们应该讲，在实际上杯

子从桌子落下并粉碎，鸡蛋的搅拌，或糖在咖啡中的溶解都是不可逆的；而少数粒子的互相反弹，还有许多能量没有损耗变成热的各种仔细控制的情形是可逆的。基本上讲，“不可逆”这一个术语只是指这样的一个事实，即不可能去追踪或控制系统中的所有个别粒子运动的所有细节。这些不可控制的运动被叫作“热”。这样，不可逆性似乎只是一个“实用的”东西。虽然按照力学定律我们完全允许去恢复鸡蛋，但在原则上这是不可能的。难道我们的熵概念要依赖于什么是可行的，什么是不可行的吗？

我们记得在第5章中，能量以及动量和角动量的物理概念可以按照粒子的位置、速度、质量和力在数学上被精确地定义。我们怎能期望“显明无序性”的概念也做到一样好，使之成为一个数学上精确的概念呢？显然，对于一个观察者“显明”并不表明对另一个观察者亦是如此。它是否取决于每位观察者对被观察系统的测量精度呢？一个观察者用一台更好的测量仪也许能比另一个观察者得到关于系统微观结构的更细致的信息。系统中更多的“隐藏的有序”也许对一个观察者是显明的，对另一个观察者却是另外一回事。相应地，前者会断言熵比后者估算的要低。不同观察者的美学判断似乎也会被牵涉到那些被定为“有序”而不是“无序”的东西。我们可以想象，有些艺术家的观点认为一堆破碎的玻璃片远比曾经待在桌子的边缘上丑陋吓人的杯子更为美丽有序！熵是否会在这种具有艺术感觉的观察者的判断那里被降低呢？

尽管存在这些主观性的问题，使人惊异的是，在精密的科学描述中熵概念是极其有用的。这一点是无疑的。这么有用的原因在于，一个系统按照细致的粒子位置和速度从有序向无序的转变是极其巨大的，并且（在几乎所有的情况下）完全把在宏观尺度上关于何为“显明有序”的观点的任何合理的差别完全淹没。特别是艺术家或科学家关于聚集或破碎的玻璃哪种更有序的判断，以熵的测度来考察，则几乎毫无结果。迄今为止对于熵的主要贡献来自于引起温度微小增加的随机的粒子运动，水的溅开以及一杯水落到地面上去等。

为了更精密地定义熵的概念，让我们回到第5章引进的相空间的观念。我们记得，系统的相空间通常具有极大的维数，其中每一点代表了包括系统的所有细节的整个物理态。相空间的一个单独的点提供了构成该物理系统的每一个单独粒子的位置和动量坐标。为了熵的概念，我们需要用一种办法把从其显明（也即宏观）性质看起来一样的所有的态集

中起来。这样，我们必须把我们的相空间分成一些区域（图7.3）。属于任何特别区域的不同点虽然代表它们粒子的位置和运动的不同细节，但是对于宏观的观察特征而言，仍然认为是一样的物理系统。从什么是显明的观点看，一个单独区域中的所有点应被考虑作相同的物理系统。相空间这样地被划分成区域的做法被称为相空间的粗粒化。

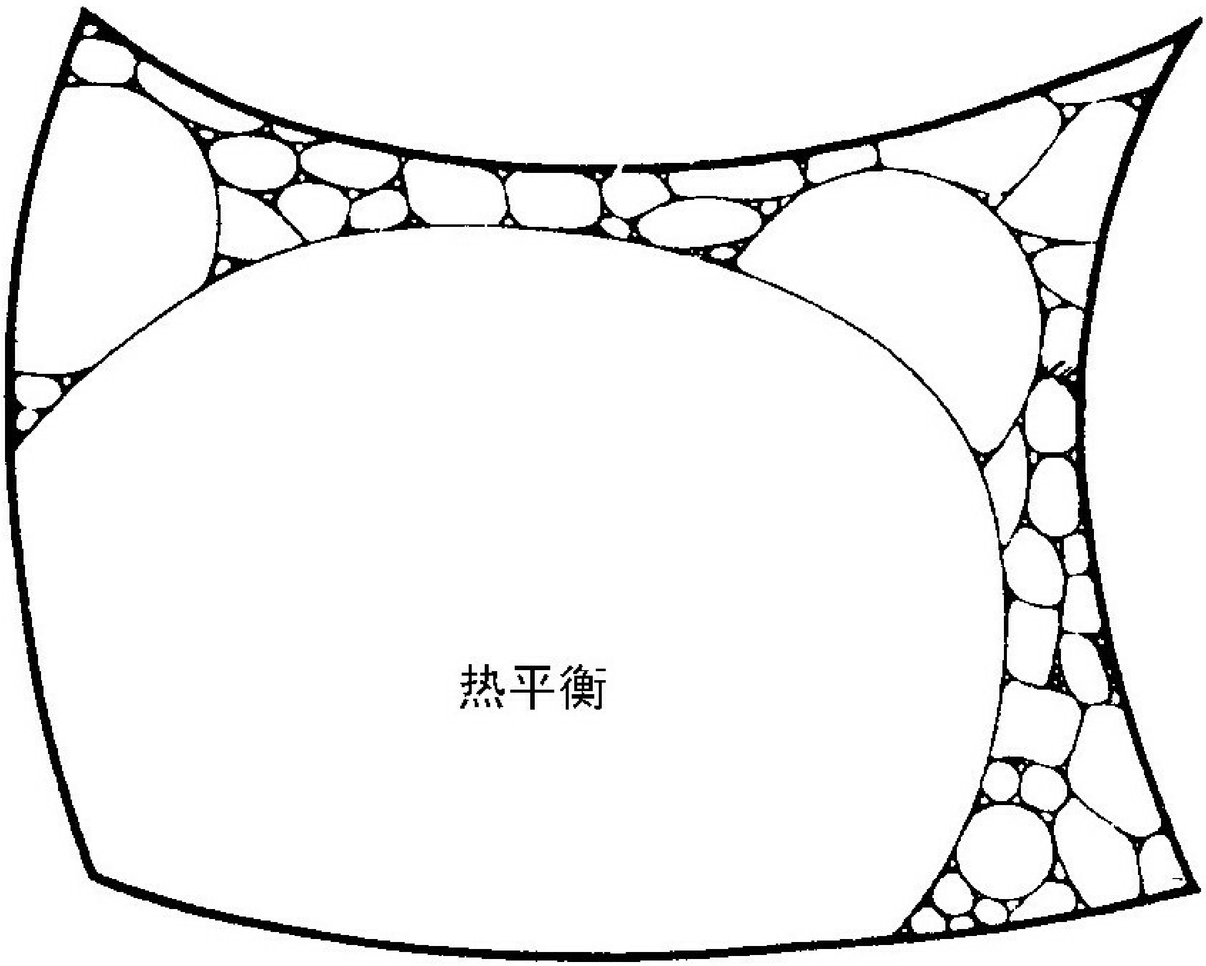


图7.3　相空间被粗粒化成在宏观上无法相互区分开的态的区域。熵和相空间体积的对数成比例

现在，这些区域中的一些会比其他的区域庞大得多。例如，考虑一盒气体的相空间。相空间的大部分体积对应于气体非常均匀地在盒子中分布的态，粒子以一种能提供均匀温度和压力的特征的方式运动。这种运动的特别方式，在某种意义上可能是称之为麦克斯韦分布的最“紊乱的”一种，它是以我们前面提到的同一位詹姆斯·克拉克·麦克斯韦来命名的：气体处于这种紊乱状态时就说它达到了热平衡。相空间中的点的绝对大的体积对应于热平衡；该体积中的点描述和热平衡一致的个别粒子

位置和速度的所有不同的细致形态。这个巨大的体积是我们在相空间中的一个（很容易是）最大的区域，实际上它几乎占据了整个相空间！让我们考虑气体的另一种可能的态，譬如所有的气体被局限在盒子的一个角落上。又存在许多不同的个别粒子的细致的态，它们都描述以同样的方式把气体局限在盒子角落的宏观态。所有这些在宏观上都不能互相区别，而相空间中代表它们的点构成了相空间的另一个区域。然而，这一个区域体积比代表热平衡的那个区域要小得多了。如果我们的盒子的体积为1立方米，装有在通常大气压和温度下的平衡的气体，而角落区域的体积取作1立方厘米，则上面的相空间体积的缩小因子大约为$10^{10^{25}}$！

为了评价这类相空间体积之间的差异，想象一种简化的情形，即把许多球分配到几个方格中去。假如每一方格或者是空的或者只容纳一个球。用球来代表气体分子而方格表示分子在盒子里所占据的位置。让我们从所有方格中挑出特殊的小子集；这些被用于代表对应于盒子的一个角落的区域的气体分子位置。为明确起见，假定刚好有1/10数目的方格为特殊的——譬如讲有n个特殊的方格和9n个非特殊的方格（图7.4）。我们希望把m个球随机地分配到这些方格中去，并且求

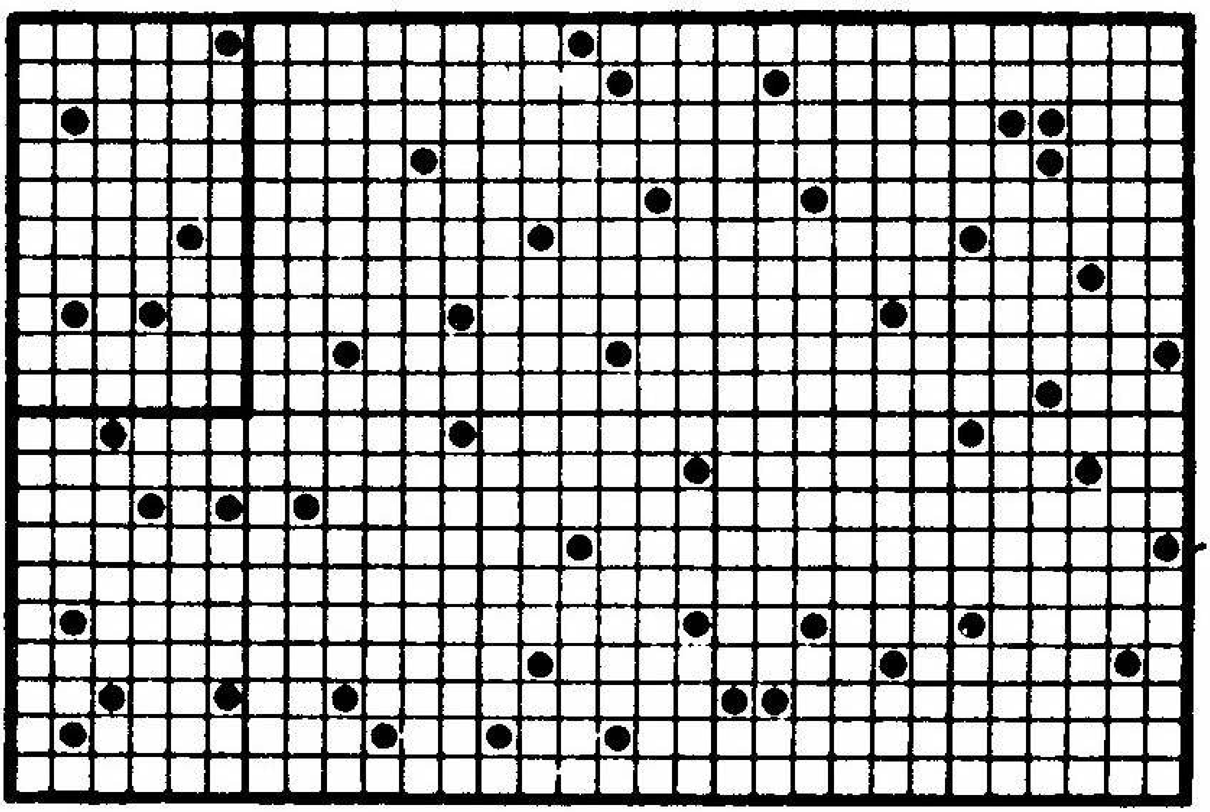


图7.4　一盒气体的模型：一些小球分布在数目比球大得多的方格中去，1/10的方格被认作特殊

的。在左上角上已把这些特殊的标出

出所有的球都落到特殊方格中去的机会。如果只有1个球和10个方格（这样我们只有1个特殊方格），则很清楚，机会应为1/10。如果只有1个球，但有任意数目10n的方格（这样我们就有n个特殊方格），则情况不变。这样就对于仅有一个原子的“气体”，把气体局限在那个角落的区域，就具有整个“相空间”体积的1/10。倘若我们增加球的数目，所有它们都在特殊方格中的机会就非常显著地减少。对于2个球，譬如讲20个方格[55]（其中2个是特殊的）（m=2，n=2），机会为1/190，或者对于100个方格（其中10个是特殊的）（m=2，n=10），机会为1/110；对于数量非常大的方格机会变成1/100。这样，对于2个原子“气体”特殊区域的体积仅为整个“相空间”的1/100。对于3球和30个方格（m=3，n=3），机会为1/4060；而对于数量非常大的方格，机会为1/1000—这样，对于3个原子“气体”特殊区域体积就为相空间体积的1/1000。对于4球和非常大量的方格，机会为$1/10^4$。对于5球和非常大量的方格，机会为$1/10^5$，等等。对于m球和大量的方格，机会为$1/10^m$。这样，对于m原子“气体”，特殊区域的体积为“相空间”的$1/10^m$（如果把“动量”也包括在内，这仍然成立。）

我们可以把这些应用于前面考虑的一盒实际气体的情形。但是现在，特殊区域不是占据总体积的$1/10^5$，而是$1/10^6$（亦即1立方米中的1立方厘米）。这表明现在的机会不是$1/10^m$，而是1/（1000000）^{m} 也就是$1/10^{6m}$。在通常的情况下，我们整个盒子中大约有10^{25}个分子，所以我们取$m=10^{25}$。这样，代表所有气体被局限在角落里的相空间的特殊区域只有整个相空间体积的

$1/10^{60000000000000000000000000}$！

状态的熵是包含代表该态的相空间区域体积V的测度。鉴于上述的这些体积间的巨大差别，最好不把它定义为和该体积成比例，而是定义为和该体积的对数成比例：

熵=klgV。

取对数有助于使这些数显得更合情理。例如10000000的对数[56]大约为16。量k称为玻耳兹曼常数。其数值大约为10^{-23}焦/开。此处取对数

的主要原因是使熵对于独立的系统成为可加量。这样，对于两个完全独立的系统，它们合并起来的系统的总熵为每一个单独系统的熵的和。这是对数函数的基本代数性质的推论：lgAB=lgA+lgB。如果系统在它们各自的相空间中属于体积为A和B的区域，则合并起来后的相空间中的区划体积就是它们的积AB，这是因为一个系统的每一可能性都必须各自分别计算。所以合并系统的熵的确为两个单独的熵的总和。

按照熵的观点，相空间中区划尺度的巨大差异显得更合理。上述的1立方米盒子的气体的熵只比集中在1立方厘米尺度的“特殊”区域的气体大1400焦/开（$=14k\times10^{25}$）［由于$\log_e(10^{6\times10^{25}})$ 14×10^{25}］。

大约为为了得到这些区划的实际的熵值，我们要稍微忧虑所选择的单位（米、焦、千克、开等）。这有点离题太远，实际上，对于我马上要给出的极其巨大的熵值，选用何种单位根本没有什么本质上的不同。然而，为了确定起见（对于专家而言），我将采用由量子力学规则所提供的自然单位，这时玻耳兹曼常数就变成1：

$k=1.$

第二定律在起作用

现在假定我们的系统从某种非常特殊的情形开始，譬如所有气体都在盒子的一个角落里。下一时刻，气体就会散开，并会急速地占领越来越大的体积。它过一阵就达到了热平衡。在相空间中看我们的图像应是什么样的呢？在每一阶段，气体所有粒子的位置和运动的完全的细节的状态都由相空间中的单独的一点描述。这一点在相空间中随着气体的演化而徘徊，这一精确的徘徊描述了气体中所有粒子的整个历史。这点从非常小的区域出发——该区域代表所有气体在盒子的一个特殊角落的所有初始态的集合。随着气体的扩散，我们运动的点进入了一个相当大的体积，这体积相应于气体以这种方式在盒子中稍微扩散开来。当气体向更远处扩散时，相空间的点继续进入越来越大的体积，新的体积以一个绝对巨大的因子使该点以前所在的体积完全相形见绌（图7.5）。在每一种情形下，一旦点进入更大的体积，（实际上）就根本没有在原先更小的体积中找到它的机会。最后它迷失在相空间中的最大的体积中——这相应于热平衡。这个体积实际上占领了整个相空间。人们可以完全放

心，我们相空间的点在真正随机的徘徊中，在任何可以想象的时刻都不可能处在更小的体积中。只要达到热平衡，无论怎么弄，这个态都好好地待在那儿。这样，我们看到了简单地表达为相空间中适当区域体积的对数测度，其系统的熵随着时间无情增加[57]的趋势。

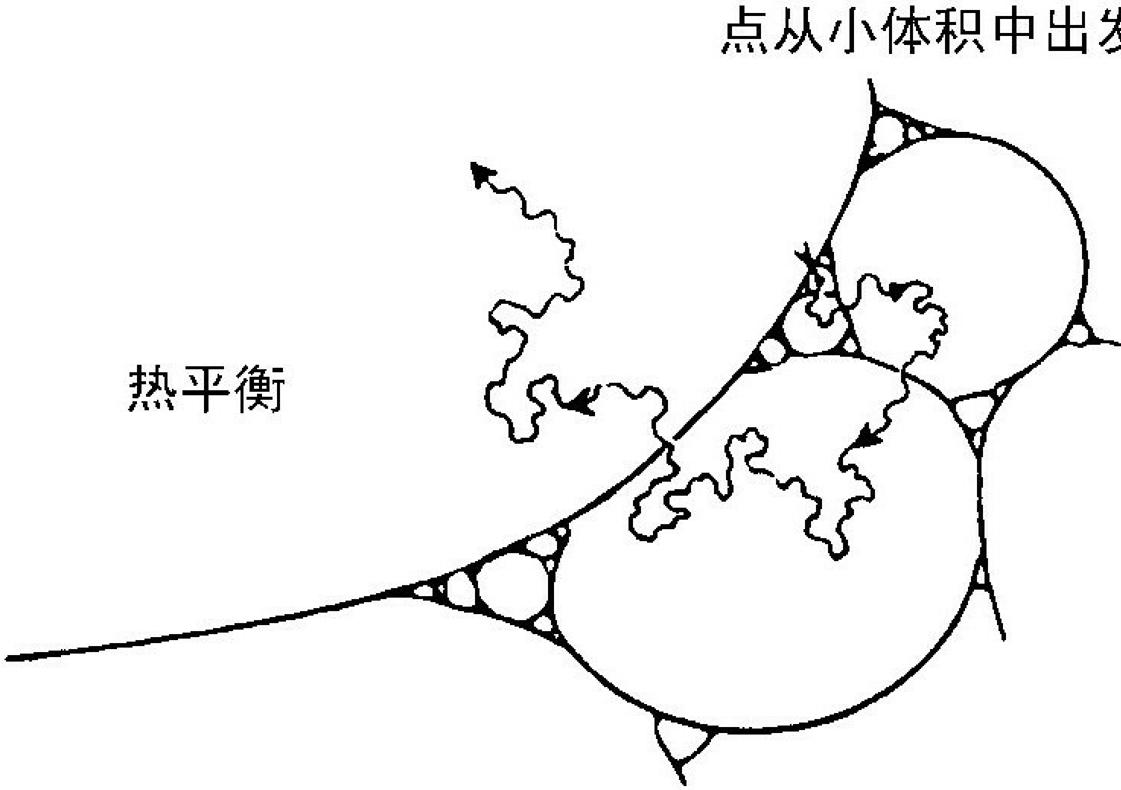


图7.5　热力学第二定律在作用：随着时间演化，相空间点进入越来越大体积的区域中。结果熵连续地增加

现在我们似乎为第二定律找到了一个解释！由于我们可以假定相空间的点不以任何特别设计的方式运动，如果它从相应于小的熵的很小的相空间体积出发，随着时间的流逝，它一定会以压倒一切的可能性不断进入越来越大的相空间体积，这相应于熵值的逐渐增加。

但是，在我们用这个论证推导出来的结果中似乎有点古怪的东西。我们似乎已经推导出时间反对称的结论。熵在时间的正方向增加，所以必须在相反的方向上减少。这个时间反对称从何而来？我们肯定没有引进过时间反对称的物理定律。时间反对称仅仅是从这一个事实而来，就是该系统从一个非常特别的（亦即低熵的）态出发，系统一旦这样地被

启动，我们就看到它在未来的方向演化并发现熵在增加。这种熵增加的确和我们自己实际宇宙中的系统行为相符。但是，我们同样可以在时间的相反方向上应用这一论断。我们又可以在某一时刻使系统处在一个低熵的状态，但是现在要问的是，什么是在此之前的最可能态的系列。

让我们试图以颠倒的方式来论证：和以前一样，从一个所有气体都待在一个角落的盒子里取其低熵态。现在相空间点处在我们以前出发的同一个微小的区域里。但是，现在让我们试着追踪它的往后方向的历史。如果我们想象，相空间中的点正如前面那样以非常紊乱的方式徘徊。随着向时间的相反方向的追踪，和前面一样地，它会很快地达到同样更大的相空间体积。这相当于气体在盒子中扩散了一些，但还没达到热平衡。体积越来越大，每一个新的体积都使原先的完全相形见绌。我们会发现，在更早的时刻它处于最大的体积中，这代表了热平衡。我们现在似乎得到推论，若在某一时刻，气体停在盒子的一个角落里，那么最可能的方式是，它是从热平衡出发才到达那里的，然后开始把自己集中在盒子的一端，最终把自己集中在盒子的一个很小的特定角落。熵在这整个过程中必须减少：它从最高的平衡值开始，然后逐渐减少，直到达到对应于气体被局限在盒子角落时的最低值！

当然，这一点也不像在我们宇宙里实际上所发生的！熵不以这种方式减少，它反而增加。如果知道在某一个特定的时刻气体挤在盒子的某一角落，那么在这之前更多得多的可能是气体被后来很快移开的一块隔板紧密地限制。或者气体以凝聚态或液态被定在该处并很快地加热成为气态。对于所有这些可能性，原先的态的熵甚至更低。第二定律的确在起支配作用，熵总在增加——也就是它实际上在时间的相反方向上减少。现在，我们看到我们的论证给出了完全错误的答案！它告诉我们使气体跑到盒子的角落去的最可能的方式是从热平衡开始，然后随着熵的逐渐减少，气体会集中到角落上去；而事实上，在实际世界中，这是极不可能发生的。在我们的世界中，气体是从一种更少可能（也即更低熵）的状态出发，挤在一个角落里的气体的熵不断增加到后来所具有的值。

我们的论证虽然不能应用于过去的方向，似乎在未来的方向上可以。对于未来的方向，我们可以正确地预料到，只要气体从角落上出发，未来最可能发生的是将要达到热平衡，而不是突然出现分隔，或气体忽然凝固或变成流体。这么奇异的可能性正是表明，我们的相空间论

证中似乎已正确地排除在未来方向熵降低的行为。但是过去的方向，这样奇异的可能性的确像是要发生似的——它们对我们而言一点也不奇异。当我们试图在相反的时间方向应用相空间论证时，我们会得到完全错误的答案！

很清楚，这给我们原先的论证投下了疑问的阴影。我们没有推导出第二定律。事实上，该论证显示的只是，对于一个给定的低熵的状态（譬如讲气体被限制在一个角落里，那么在不存在任何约束此系统的外在因素时，则可望熵从该给定的状态在时间的两个方向上增加（图7.6）。这个论证在时间的过去方向上无效正是因为存在这种因素。过去的确有某种东西在约束这个系统。某种东西强迫熵在过去取低的值。熵在将来增加的这种趋势不足为奇。在某种意义上讲，高熵的态就是自然的“态”，这点就不必多加解释了，但在过去的低熵态是令人困惑的。是什么约束使得我们世界的过去的熵变得这么低？具有令人不可思议的低熵状态在我们居住的实在宇宙中普遍存在，虽然我们对这一点早已司空见惯，并通常不认为有什么大惊小怪，但它的确是一个令人惊异的事实。我们自己本身便是具有极小熵值的结构！从上述的论证可以看出，给定一个低熵态，我们不应该为后来的熵增加感到惊讶。应该惊讶的是，当我们考察它的过去时，熵变得越来越不可想象的低！

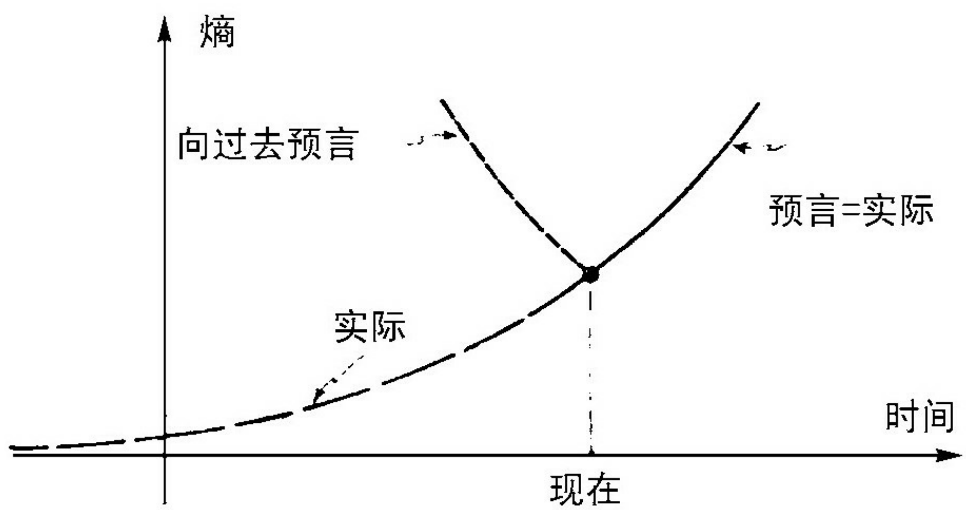


图7.6　如果我们在时间的颠倒方向上应用画在图7.5的论证，我们就“向过去预言”熵从它现在的值也向过去的方向增加，这和观察严重冲突

宇宙中低熵的起源

我们将要理解在我们居住着的现实世界中“惊人的”低熵从何而来。让我们从自身开始。如果我们能理解我们自身的低熵从何而来，则我们就应能看到被隔板限制住的气体、桌子上的水杯、炒锅上的鸡蛋或悬在一杯咖啡上的糖块的低熵从何而来。一个人或一群人（或者一只母鸡！）直接或间接地为每一种情形负责。我们自身的一部分低熵实际上有很大的程度被用以建立这其他的低熵态。也许牵涉一些附加的因素，例如使用真空泵把气体注入到隔板后面去。如果这台泵不是人工驱动的，则必须用某种“化石燃料”（例如石油）燃烧以提供必要的低熵能量使之运转。也许这台泵是电动的，则在一定的程度上要依赖于贮藏在核电站的铀燃料的低熵能量。以后我还会讲到其他低熵的源，但是现在我们先考虑自己身上的低熵。

我们自身的低熵究竟从何而来呢？我们身体的组织是由我们吃的食物和我们呼吸的氧气来的。人们经常听到这样的说法，即我们从食物和氧气的摄入中得到能量。但是只要想得更清晰一些就会发现这不是完全正确的。的确，我们消耗的食物和吸收到身体中来的氧气的化合为我们提供了能量。但是，大多数情况下，该能量又重新以热的形式离开我们的身体。由于能量是守恒的，在我们整个成年的生活中，身体实际上的能量含量或多或少是维持着一个常量，我们身体一点也没有必要再添加能量。我们不需要比我们已具有的更多的能量。事实上，当我们的体重增加时我们的确添加了能量，但通常这是多余的！还有，当我们从儿童长大，体格变健壮时，能量含量增加了相当多；这不是我在这里所关心的。问题在于我们如何使自己在正常（主要成年的）生活中存活。我们不必为此增加自身的能量含量。

然而，我们确需要取代以热的形式连续损失的能量。事实上，我们越是“有精力”，则实际上以这种方式损失的能量越多。所以这能量都必须有所取代。热量是能量的最无序的形式，也就是说，它是能量的最高熵的形式。我们吸收低熵形式的能量（食物和氧气）并以高熵形式（热、二氧化碳、排泄物）排泄出去。我们没必要从我们的环境获取能量，因为能量是守恒的。但是，我们是在连续地对抗热力学第二定律。熵不守恒，它无时无刻地增加着。我们必须使自身的熵降低才能存活。为此我们从食物和大气氧气中吸收低熵的化合物，让它们在我们身体内化合，以高熵的形式释放能量，否则我们的体重就会增加。用这种方

式，我们可维持我们身体内的熵不增加，并能保持（并甚至增加）我们的内部组织。（见*Schrödinger 1967* 。）

从什么地方来提供这些低熵呢？如果我们吃的食物刚好是肉（或蘑菇），那它正如我们一样要依赖于更外部的低熵源去提供和维持其低熵结构。这只不过把我们外部的低熵源的问题推到其他的地方。这样，让我们假定我们（或动物或蘑菇）消化植物。我们因为绿色植物的巧妙——不管是直接的或是间接的——而必须极其感谢它：因它吸收大气中的二氧化碳，把氧气从碳中分离开来，而利用碳来建造它们自身的结构。这一光合作用的过程导致大量的熵降低。我们自己实际上在身体内把氧和碳重新简单地结合，用这种办法利用低熵的这种分离。绿色植物为什么能实现熵降的魔术呢？它们是利用阳光来实现的。阳光给地球带来了相当低熵形式的能量，即是可见光光子的能量。地球，包括它上面的居住者，不能保留此能量，而是（过了一阵）就把它全部重新辐射回到太空去。然而重新辐射的能量具有高熵的形式并被称为“辐射热”——它表明是红外光子。和普遍的印象正相反，地球（和居住者）并不从太阳获得能量！地球所进行的只不过是取来低熵形式的能量，然后以高熵的形式全部把它吐回到太空去（图7.7）。太阳对我们所做的是给我们提供了巨大的低熵源。我们（通过植物的巧妙功能）利用了这些低熵，最终抽取某一极小的部分将其转换成惊人的、错综复杂的、有组织的结构，这就是我们自身。

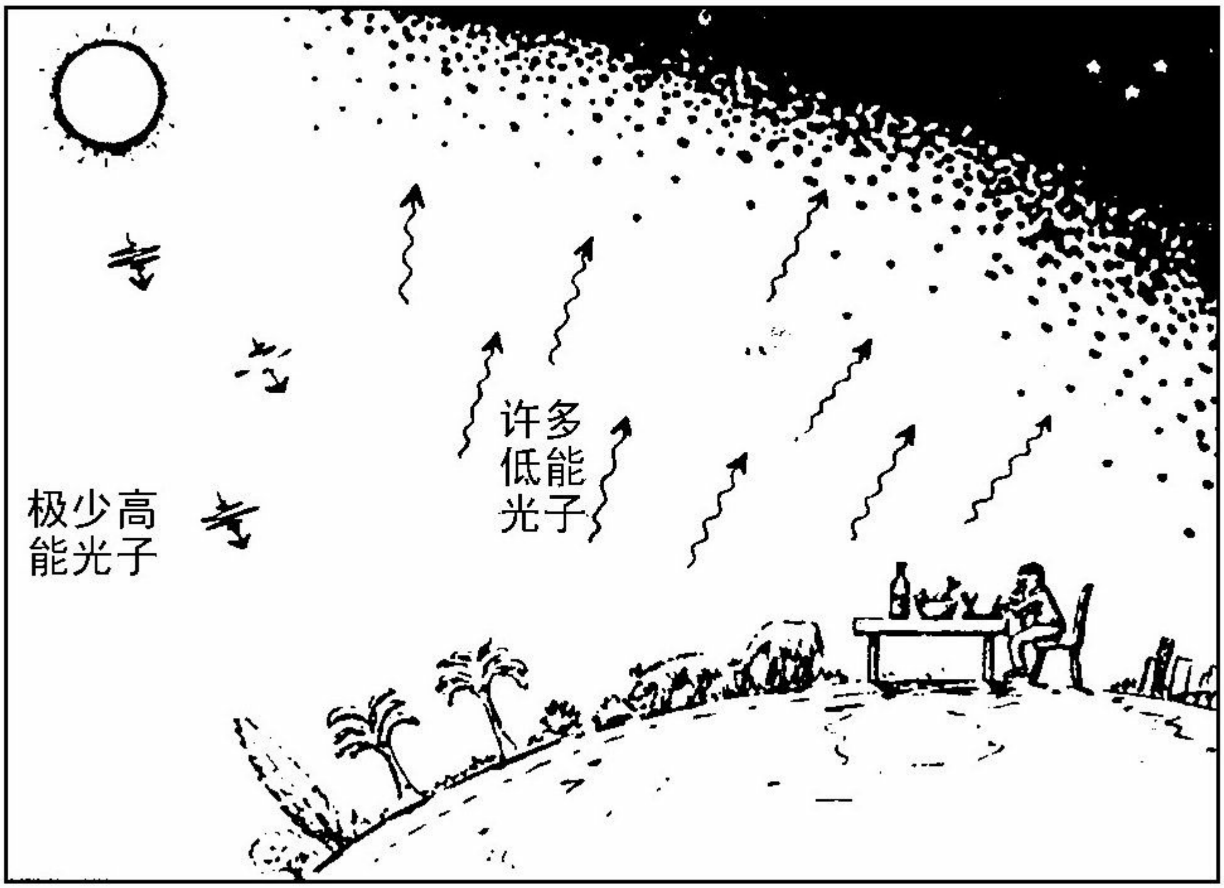


图7.7　我们如此利用这事实：太阳是黑暗太空中的一个热点

让我们以整体的观点考虑太阳和地球，能量和熵发生了什么变化？太阳以可见光光子的形式辐射能量。其中有一些被地球所吸收，它们的能量以红外光子的形式被重新辐射。现在，在可见光和红外光之间的关键差别在于前者有一个高频，所以单独光子比后者有更高的能量。（回忆一下在292页给出的普朗克公式E=hν。它告诉我们，光子的频率越高则能量越大。）由于每一可见光子比每一红外光子具有更高的能量，为了使进入地球的和离开地球的能量相平衡，只能有比离开地球的红外光子的数目更少的可见光子到达地球。地球吐回到太空去的能量被分散到比从太阳接收到的能量的多得多的自由度去。由于把能量再送回太空时牵涉到多了这么多的自由度，相空间的体积变得大得多，所以熵值就被极大地增加。绿色植物吸收低熵形式的能量（相对少量的可见光子）而重新把它以高熵形式（相对多量的红外光子）辐射，为我们提供了所需要的分解的氧和碳，以这种方法把低熵喂给我们。

所有这一切之所以可能的原因是，太阳为天空中的一个热点！天空

处于温度不平衡的状态；它的一个小区域亦即太阳占据的地方，比其他地方的温度高得多。这个事实为我们提供了所需要的强大的低熵源。地球从这一个热点得到低熵形式（少量光子）的能量，然后以高熵的形式（许多光子）重新辐射到冷的区域去。

太阳为什么是这样的一个热点呢？如何才能得到这个温度的不平衡，并因此为我们提供低熵态呢？答案是它从原先均匀分布的气体（主要是氢气）的引力收缩形成的。在其形成的早期阶段，当它收缩时，太阳被加热上去。在到它的温度和压力达到一定点之前，也即除了引力收缩外，它还找到另一种叫作热核反应的能源，它会继续收缩并变得更热。热核反应使氢核聚变成氦核，并同时释放能量。如果没有热核反应，太阳会变得比目前的更热得多和小得多，直到最终消逝。热核反应使太阳不再继续收缩以免过热，从而使它稳定在适合于我们的温度上，能在更长久的时间里持续发光，否则的话早已熄灭。

意识到这一点是很重要的。虽然热核反应在决定从太阳辐射来的能量的性质和多少方面无疑极具意义，但是引力才是关键之所在。（事实上，热核反应的潜力对太阳的低熵值的确有很大的贡献，但是聚变的熵引起了微妙的问题，更充分的讨论，只使论证更为复杂，而不影响最终的结论。）[3] 没有引力，甚至太阳根本就不会存在！没有热核反应太阳仍然发光——虽然不以适合我们的方式——但是没有引力就根本没有发光的太阳，的确需要引力来聚合物质，并提供所需要的温度和压力。若无引力，代替太阳之处我们只会有一团冷而弥散的气体，在天空中不会有热点！

我未讨论到地球内“化石燃料”中的低熵来源，但是其考虑基本上是一样的。根据传统理论，地球上所有的油（和天然气）是来自于史前植物的生命，又是植物被当作低熵的来源。这些史前植物从太阳得到它们的低熵——所以我们应该再次转向把弥散气体变成太阳的引力作用。托玛斯·高尔德提出了地球上石油起源的离经叛道的理论。他不同意传统的观点，认为地球上存在比史前植物产生的更丰富得多的碳氢化合物。高尔德认为，油和天然气是在地球形成时被包含在地球内的，并一直连续地渗透出来直到下层的矿穴 [4] 。根据高尔德的理论，油在地球形成之前，即使在外空仍然也是由阳光合成的。这又是起源于引力形成的太阳。

用于核电站的铀235同位素的低熵核能量又如何呢？这的确不是原先从太阳（虽然在某阶段它也可能通过太阳），而是从某些其他的恒星来的。这些恒星在几十亿年前的一次超新星爆发中爆炸！这些物质实际上是从许多这类爆炸的恒星中聚集起来的。爆发把这些物质从恒星中吐到太空去，其中一些最终（通过太阳的作用）被聚集在一块，并把重元素提供给地球，包括它所有的铀235。每一个核子以及其低熵能量的储藏是来自于发生在某次超新星爆发的激烈的核过程中。这种爆发是发生于恒星的引力坍缩 [5] 的余波。当恒星的质量过大，以至于热压力不能支持其自身时就会坍缩。一个小的核——可能以所谓的中子星的形式（后面还要更详细地讨论）在坍缩和紧随着的爆发之后残存下来。恒星原先是从弥散的气体云收缩而来，包括我们的铀235的许多原始物质又都被抛回到太空中去。残留下的中子星从引力收缩中得到了巨大的熵。引力再次成为最主要原因——这一次它把弥散的气体凝聚成（过程最终是激烈的）一个中子星。

我们似乎得出这样的结论，第二定律中最令人困惑的方面即所有在我们四周发现的明显的低熵，应归结于这样的一个事实，即通过弥散气体引力收缩成恒星的过程中可得到大量的熵。所有这些弥散气体从何而来？这些气体从弥散状态开始的这一事实为我们提供了大量的低熵储藏。我们正在消耗这种低熵的储藏，并将在未来的漫长岁月里继续如此。正是这些气体引力结团的潜力给我们带来了第二定律。此外，不仅仅是引力结团产生的第二定律，而且还有比下面简单陈述更精密和细致得多的某种东西："世界是从非常低的熵开始的。"我们还可以用其他不同的方式得到"低"的熵，也就是说在早期的宇宙中有巨大的"显明有序"，但是这和在实际上呈现给我们的"有序"完全不同。（想象早期宇宙也许是正规的十二面体——这或许会投合柏拉图的心意——或者是其他某种不像会发生的几何形状。这的确是"显明有序的"，但并非我们预期在实际的早期宇宙中所发现的那种形状！）我们必须理解所有这些弥散气体从何而来——为此，我们必须转向宇宙论的研究。

宇宙论和大爆炸

我们如果使用最强大的望远镜——不管是光学的还是射电的，就会发现宇宙在非常大的尺度下显得相当均匀；但是更惊人的事实是，它正在膨胀。我们观测得越远，则遥远星系（以及甚至更远的类星体）就显

得越快速地从我们这里离开。似乎宇宙本身是从一个巨大的爆炸事件中产生——这一个事件称为大爆炸，它发生在大约100亿年以前[58]。所谓的黑体背景辐射对于宇宙的均匀性以及大爆炸的实际存在提供了印象深刻的支持。它就是一种光的杂乱运动，而且是分辨不出来源的热辐射——其绝对温度大约为2.7度（2.7开），也就是-270.3摄氏度和-454.5华氏度。这似乎是非常冷的温度——也的确如此——但是它乃是大爆炸本身的那一瞬间的残留！因为从大爆炸的时刻以来，宇宙膨胀了这么巨大的因子，原始火球以一绝对巨大的因子发散开来。大爆炸的温度远远超过现在所能发生的温度，但是由于膨胀，该温度被冷却到今天微波背景所具有的微小的数值。1948年，美籍苏联物理学家和天文学家乔治·伽莫夫用现今标准的大爆炸图像作基础，预言了这个背景的存在。在1965年，彭齐亚斯和威尔逊首次（意外地）观测到它。

我应该阐释经常给人们带来困惑的一个问题。如果宇宙中所有的远处星系都离开我们而去，是不是意味着我们自身在宇宙中占据着某种非常特别的中心位置呢？不，不是这样！不管我们位于宇宙中的何处，都会看到远处星系的同样的退缩。该膨胀在大尺度上是均匀的，没有一个位置比其他的更优越。通常可以用被吹胀的气球来描绘这种情景（图7.8）。假定在气球上存在代表不同星系的斑点，取气球本身的二维表面代表整个三维类空的宇宙。可以清楚地看到，所有气球上其他的点都从气球上的每一点退走。在这个方面，气球上没有一点比其他点更优越。类似地，从宇宙中的任何一个星系的有利地点来看，所有其他的星系在任何方向都同等地从它那里退走。

三种标准的所谓弗里德曼——罗伯逊——瓦尔克（FRW）宇宙模型之一，即空间封闭的正曲率的FRW模型，膨胀气球提供了非常好的图解。在另外两种（零或负曲率的）FRW模型中，宇宙以同样方式膨胀，但是这回空间不像用气球表面上标出的有限宇宙，我们拥有包含了无限数目星系的无限的宇宙。

这两种无限模型中的较简单的一种是空间几何为欧几里得的那种，也就是具有零曲率的。用一个通常的平面代表整个空间的宇宙，画在上面的点代表星系。当宇宙随着时间演化，这些星系以一种均匀的方式相互离开。让我们按照空间——时间来考虑。我们对每一“时刻”都有一个相应的而且不同的欧几里得平面，把这些平面想象成一个重叠在另一个上面。这样，我们一下子就有了整个空间——时间的图像（图7.9）。

现在星系可用曲线——也就是星系历史的世界线——来代表，它们在未来的方向上相互离开。没有任何星系的世界线是特别的。

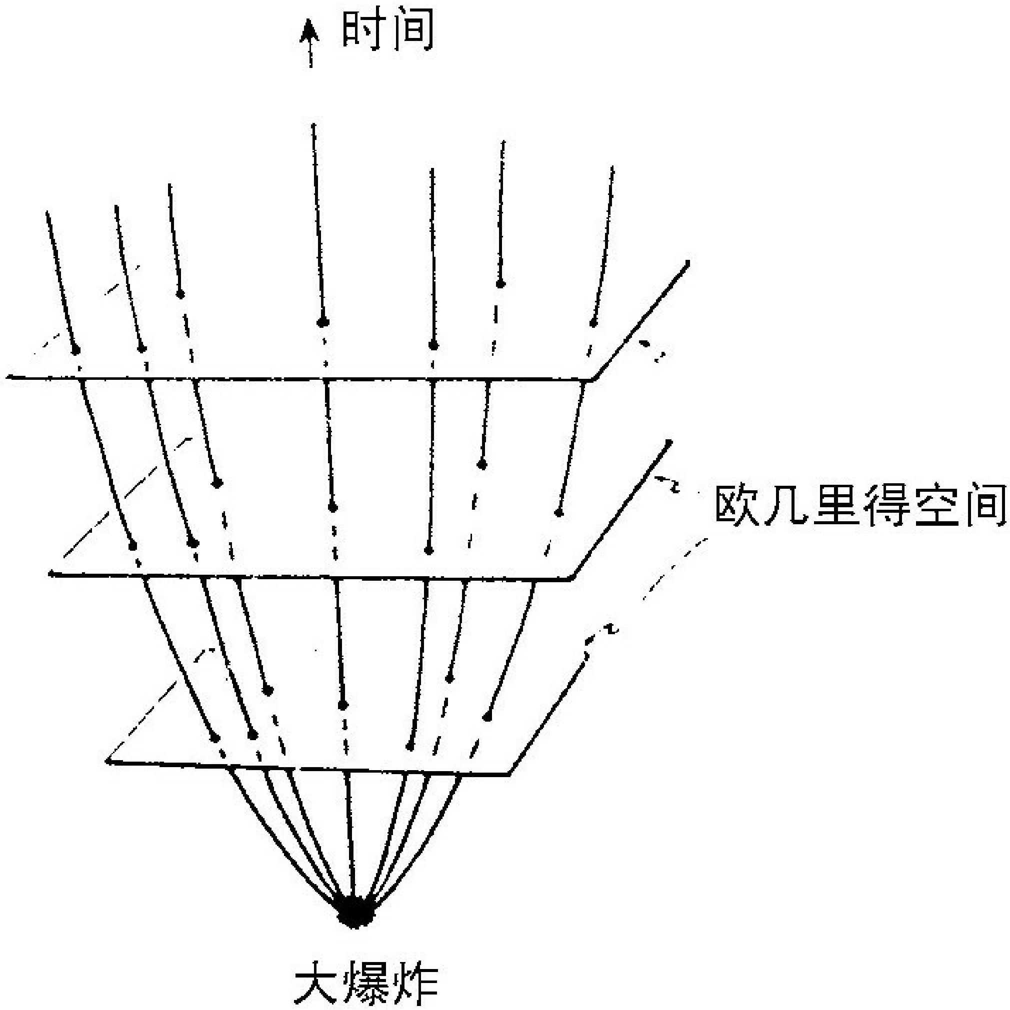


图7.9　具有欧几里得空间截面的膨胀宇宙的空间——时间图（画出了空间的两维）

对于星系的另一种FRW模型，也就是负曲率的模型，空间几何为非欧几里得的罗巴切夫斯基几何，这种几何已在第5章中描述过并用图5.2（204页）的埃索图来解释。在空间——时间描述中，我们在每一“时刻”都需要一个罗巴切夫斯基空间，我们并把这些一个重叠一个以构成整个空间——时间的图（图7.10）[6]。星系的世界线又是在未来方向相互离开的世界线，没有什么星系是特别选择的。

当然，在我们所有的这些描述中，空间的三个维中有一个被压缩掉

了（正如我在第5章所做的，参阅250页），其目的在于给出比万不得已必需的完全的四维空间——时间图更易摹想的三维空间——时间图。甚至到了这种地步，如果不抛弃另一空间的维去摹想正曲率的空间——时间仍然非常困难！让我们就这么做，用一个（一维）圆周来代表正曲率的闭合的类空宇宙，而不用作为气球表面的（二维）球面。当宇宙膨胀时，这些圆圈的尺度变大。我们可把这些圆周（每一圆周代表一个“时刻”）一个一个地叠起来，结果得到一种弯曲的锥[图7.11（a）]。现在，从爱因斯坦的广义相对论方程得出，这种正曲率的闭合的宇宙不能永远地继续膨胀下去。在它达到最大尺度的阶段后，就会坍缩回去，最后会在一种倒转的大爆炸中达到零尺度[图7.11（b）]。有时把这种时间倒转的大爆炸称作大挤压。负曲率和零曲率（无限的）宇宙的FRW模型不会以这种方式坍缩。它们不会导致大挤压，而是继续无限地膨胀下去。

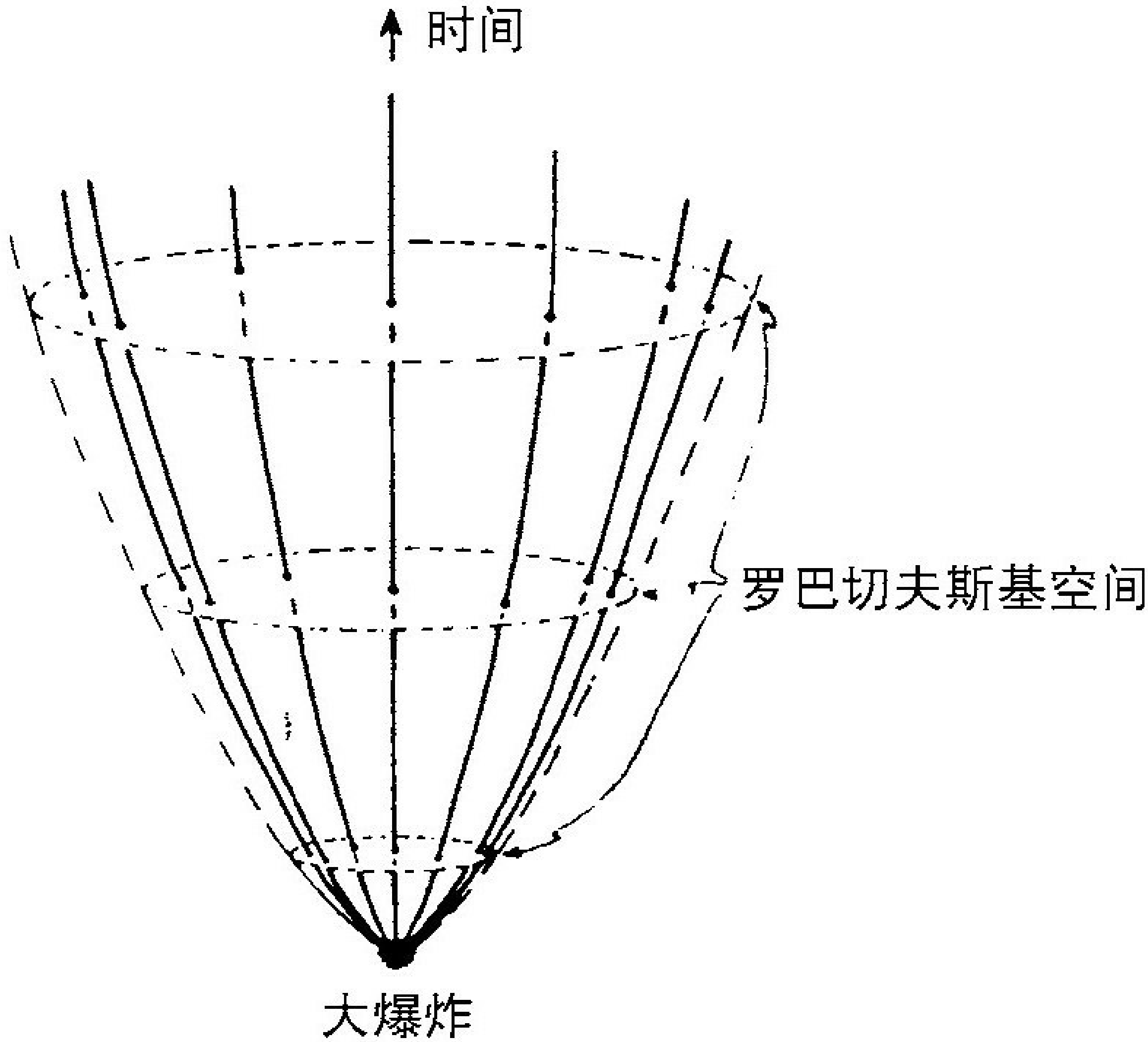


图7.10　具有罗巴切夫斯基空间截面的膨胀宇宙的空间——时间图（画出了空间的两维）

至少在所谓宇宙常数为零的标准的广义相对论中，这是对的。具有适当的非零的宇宙常数，空间无限的宇宙有可能会坍缩成大挤压，或者有限的正曲率的模型会无限地膨胀下去：非零宇宙常数的存在会使这些讨论变得稍微复杂一些，但是对于我们的目的不会有任何重大的影响，为了简单起见，我把宇宙常数取为零 [59] 。在写此书之际，从观测上知道宇宙常数是非常小的，其数据与零是一致的。（为了对宇宙模型有更多了解，可参考*Rindler 1977* 。）

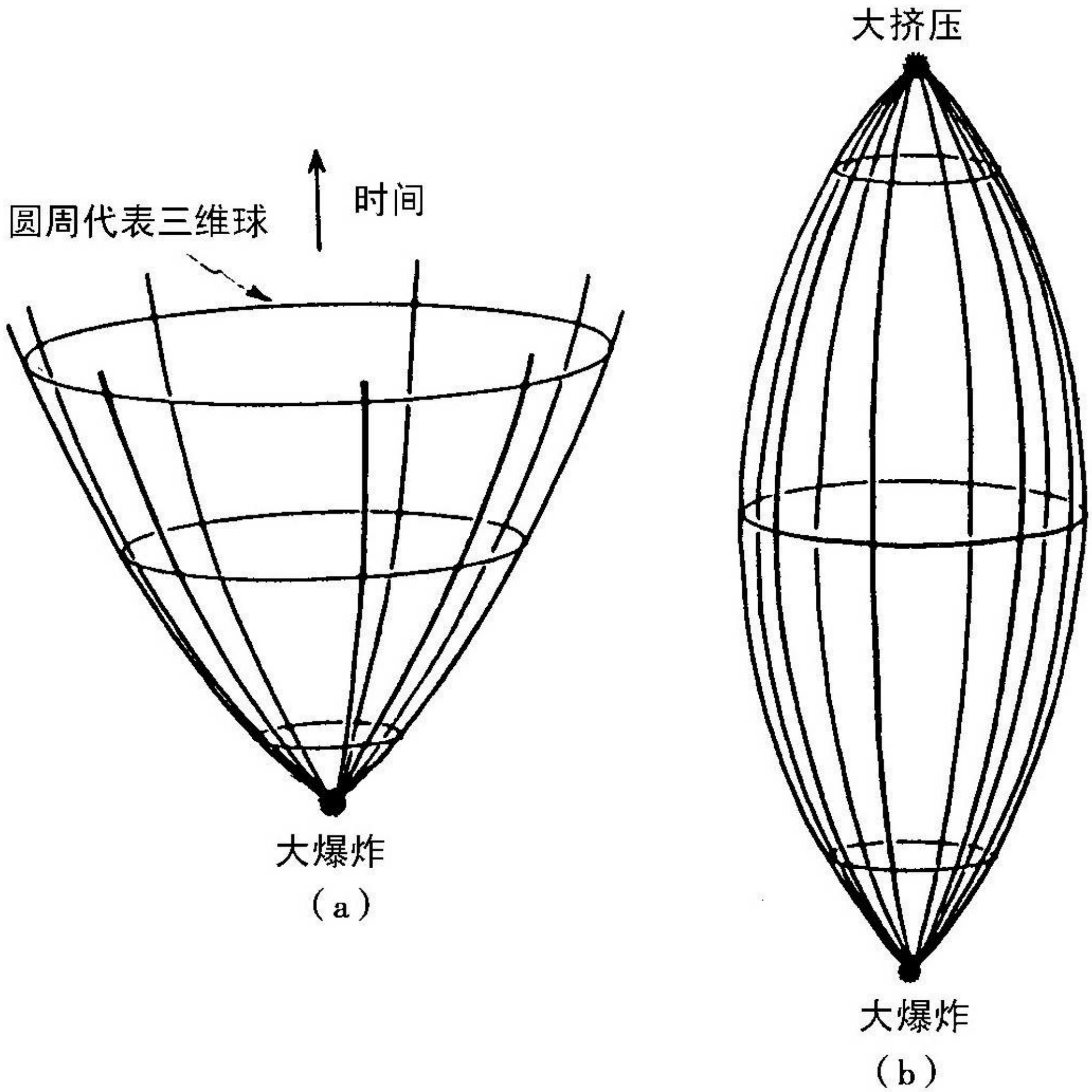


图7.11（a）具有球形空间截面（只有空间的一维被画出来）的膨胀宇宙的空间——时间图（b）这个宇宙最终会坍缩成最后的大挤压

不幸的是，我们的数据还没好到足以清楚指出，我们的宇宙应是哪一种模型（也不能确定是否存在很小的宇宙常数，可以有重大的整体的效应）。表面上看来，数据似乎表明宇宙是类空地负曲率的（在大尺度上为罗巴切夫斯基几何），而且它会继续永远地膨胀下去：这主要是基于似乎以可见形式呈现的实际物质总量的观测。然而，也可能有大量的不可见物质散布在整个太空中。宇宙在这种情形下可以是正曲率的，并可能最终坍缩到大挤压去——虽然只会在大约10^{10} 年，也就是比宇宙已经存在的这么长的时间更长得多的时间尺度下发生。要使这种坍缩发

生，必须存在大约为用望远镜可直接辨别的物质的30倍的被假想地称为“暗物质”的，充满太空的不可见物质。的确有好些间接证据表明大量暗物质的存在，但是否足够“去封闭宇宙”（或使空间平坦）——并且坍缩——还在未定之天。

太初火球

让我们回到寻求热力学第二定律起源的问题上来。我们已经把它追踪到恒星由其凝聚而成的弥散气体的存在。那么气体又是从何而来的呢？它主要是氢，但仍有大约23%（按质量计算）的氦和少量其他物质。根据标准理论，这气体是由创造宇宙的爆炸——大爆炸吐出来的结果。然而，很重要的一点是，这不是我们通常熟悉的爆炸。在那里，物质从某一个中心点喷射到一个预先存在的空间中去。而在这里，空间本身由此爆炸创生出来，并从来不存在任何中心点！这种情形也许在正曲率模型中最容易摹想了。重新考虑图7.11或者图7.8中的吹胀的气球。并不存在任何大爆炸产生的物质可注入的“预先存在的空虚空间”。空间本身也就是“气球表面”是由爆炸产生的。我们必须意识到，为了摹想的方便，在正曲率模型的图像中利用了一个“包容空间”——也即气球所在的欧几里得空间——这个包容空间并没有任何物理实在性。在气球的内部和外部的空间只是用来帮助我们摹想气球的表面。只有气球表面本身才代表了宇宙的物理空间。现在我们看到了，并不存在一个让大爆炸产生的物质从该处发散出来的中心。刚好在气球中心的点不是宇宙的部分，而仅仅是用来帮助我们去摹想这一模型。大爆炸喷出的物质均匀地发散到整个宇宙的空间！

其他两种标准模型的情形也是一样的（虽然要摹想它们更困难一些）。物质从未集中于空间中的任何单独的一点。它从一开始就充满了空间的全部！

这个图像是称为标准模型的热大爆炸理论的基础。按照这种理论，宇宙在其产生后的一瞬间处于极热的，称作太初火球的状态。关于这个火球的性质和成分以及当这火球（整个宇宙）膨胀并冷却时，这些成分如何变化都进行了细致的计算。对于描述宇宙的和我们现在如此不同的状态所进行计算的可靠性真是令人印象深刻！然而，只要我们不过问在创生后10-4秒以前发生什么的话，作为计算基础的物理学是无可争议

的！从那个时刻也就是创生后的1/10000秒后，直到后来的3分钟，宇宙的行为已被非常仔细地算出（参阅Weinberg 1977 ）——而且奇异的是，我们从现在处于非常不同状态的宇宙的实验知识推导而来并很好建立的物理理论，对于这种计算是完全足够的 [7] 。这些计算的最后结论是，许多光子（也就是光）、电子和质子（氢的两种成分）、一些α粒子（氦的核），还有少量重氢核（一种氢的同位素）和其他种类核的踪迹，也许还有大量的诸如中微子等的，几乎其存在不能被觉察得到的“不可见”粒子，都以一种均匀的方式散布在整个宇宙。其物质的成分（主要是质子和电子）会结合在一起，产生了恒星（主要是氢）在大爆炸后大约10^8 年由之形成的气体。

然而，不会立即形成恒星。在气体的进一步膨胀和冷却之后，为了局部的引力效应能开始战胜全局膨胀，某些区域的气体的相对集中是必需的。我们在这里进入了尚未解决且富有争议的星系实际上是如何形成的，以及星系可能形成的必需的初始无规性应是什么样子的问题。我不想对这些问题进行争论。我们只要接受，在初始气体云中应该存在某种无规性，引起了引力结团的某种正确方式，从而形成了包括千亿个恒星的星系。

我们已经找到弥散气体从何而来。它是从大爆炸本身的那个火球而来。正是该气体被极其均匀地分布于整个太空的事实带来了第二定律，在引力结团使熵增加的过程成为可能之后，我们就晓得了这定律的细节。实际宇宙中的物质是怎样均匀地分布呢？我们注意到恒星聚集在一起形成星系；而星系聚集在一起形成星系团；星系团组成所谓的超星系团，甚至还有某些证据，这些超星系团聚集成更大的称为超星系团集合体的集团。然而，重要的是要注意到，所有这些无规性以及结团和整个宇宙结构的令人印象深刻的均匀性相比较都是“微小的”。能够往回观测的时间越早，则宇宙被测量的部分就越大，宇宙就显得越一致。黑体背景辐射为此提供了最令人印象深刻的证据。它特别告诉我们，当宇宙年龄仅仅为100万年时，在现在已扩展开到10^{23} 千米的范围内——这是一个从我们这里开始能包含10^{10} 个星系的距离——宇宙和它的所有的物质内容都均匀到1/100000（参阅*Davies et al.1987* ）。尽管宇宙的起源是非常激烈的，它在早期的确是非常一致的。

这样，正是这个太初火球把这气体在整个太空发散得如此均匀。我们的探索也就是从此处开始。

大爆炸能解释第二定律吗

我们的探索到达尾声了吗？是否仅仅由宇宙是从大爆炸开始的情景，就能解释在宇宙中熵的初始值是如此之低，并因此导致热力学第二定律的令人困惑的事实？稍微想一下就会发现这个观念中有一些矛盾。它不能是真正的答案。回想一下太初火球是一种热的状态——处于膨胀的热平衡的热气体。还有术语“平衡”是指具有最大熵的状态。（这就是我们在提到一盒气体的最大熵状态时说到的。）然而，第二定律要求，我们宇宙的熵在其初始态处于某种极小，而不是极大！

何处出了毛病？一个“标准的”答案应该大体上如下：

是的，火球在刚开始时实际上是处于热平衡，但是那个时刻的宇宙非常微小。火球所代表的是那一微小尺度的宇宙所能允许的最大熵的状态，但是这种允许的熵和在今天宇宙尺度下能允许的熵相比较是微不足道的。随着宇宙膨胀，可允许的最大熵随着宇宙尺度增加，但是宇宙中的实际的熵远远落在允许的最大值后面。由于实际的熵总是拼命去追赶允许的最大值，所以产生了第二定律。

然而，稍微考虑一下便知道，这不应该是正确的解释。如果真是如此，在一个最终坍缩到大挤压的空间闭合的宇宙模型中，该论证在时间的颠倒方向上最终又能适用。适合于膨胀宇宙极早期并给予了我们低熵的同一限制应该又能适用于收缩宇宙的最后阶段。“时间开端”处的熵限制给了我们第二定律。根据第二定律，宇宙的熵随时间增加。如果把同一低熵的限制应用于时间的终结处，则我们应该在那里发现和热力学第二定律的严重冲突！

当然，我们实在的宇宙也许永远不会以这种方式坍缩。我们也许生活在零曲率（欧几里得情形）或负曲率（罗巴切夫斯基情形）的宇宙中。我们也许生活在一个（正曲率）坍缩的宇宙中，但是坍缩将在这么遥远的时刻发生，现在我们觉察不到对第二定律的任何违反——尽管从这种观点看，宇宙的总熵会倒转并减小到微小的值，从而按我们今天的理解，第二定律会被严重地违反。

实际上，我们有非常好的理由怀疑，在一个坍缩的宇宙会有这种熵的反转。其中最有力的原因必须和称为黑洞的神秘物体相关。在一个黑

洞中有一个坍缩宇宙的微宇宙；这样，如果在坍缩中熵的确要倒转，那么在一个黑洞附近必须能观察到第二定律的严重违反。然而，所有理由都使人相信第二定律强有力地支配着黑洞。黑洞理论为我们的熵的讨论提供了生动的内容，所以我们有必要稍微仔细地考虑这些奇怪的物体。

黑洞

让我们首先考虑理论所预言的关于我们太阳的最终命运。太阳已经存在了大约50亿年。它再过50或60亿年就会在尺度上开始膨胀，它会无情地肿大，直到表面大致达到地球的轨道。那时它就变成为称作红巨星的一种恒星类型。在天空中的其他地方能看到许多红巨星，两个最著名的是在金牛座的阿尔德巴伦和猎户座的贝特勒宙斯。在其表面膨胀的全过程中，在它的核心会有一个异常紧密的物质浓缩体，在逐渐地变大。这个紧密的核心具有白矮星的性质（图7.12）。

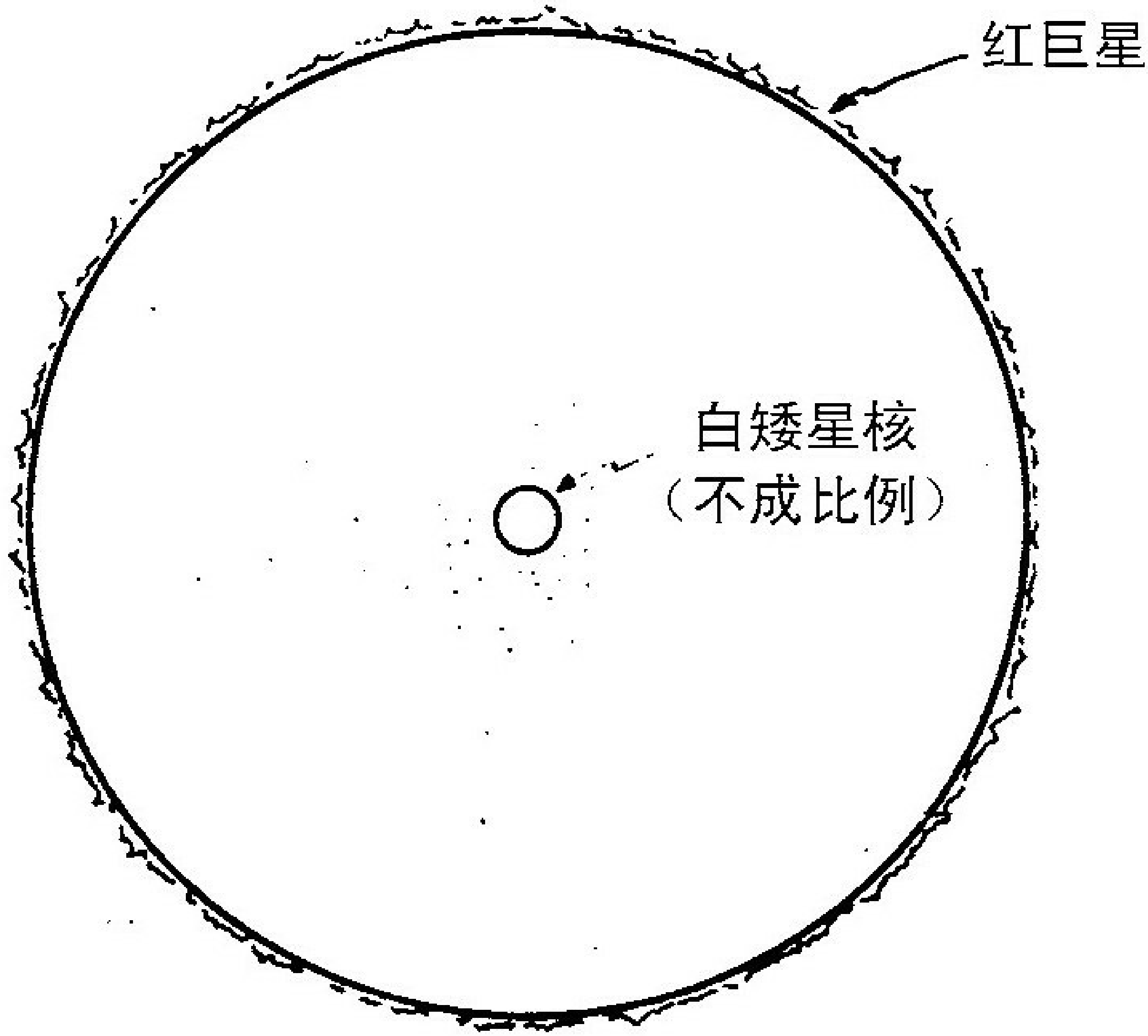


图7.12　一个红巨星及其白矮星核心

白矮星自身实际上是物质集中到极高密度的恒星。它的密度相当于一个乒乓球的体积充满了几百吨重的物质！在天空中可以观察到相当数目的这类恒星：也许在我们银河系发现的恒星中有百分之十几为白矮星。最著名的白矮星是天狼星的伴星，其惊人的高密度在本世纪初给天文学家带来了巨大的观察上的困扰。然而，后来这同一颗恒星为（在1926年左右R.H.否勒开创的）物理理论提供了美妙的证实。根据这个理论，有些恒星的确可以具有这样巨大的密度，该恒星由“电子简并压力”支撑着。这表明当泡利量子力学的不相容原理（参阅第354页）应用于电子时，可以防止恒星遭受向内的引力坍缩。

任何红巨星的核心都有一个白矮星，这个核心会继续从恒星的主体收集物质。红巨星最终被这个寄生的核完全消耗，而大约有地球那样尺度的实际的白矮星成为仅有的幸存者。可以预料，我们的太阳作为红巨星的形式“仅仅”会存在几十亿年。然后，在它的最后的“可见”肉身——作为一个慢慢冷却地死去的白矮星的余烬 [60] —太阳将再维持几十亿年，最后完全变阴暗了，成了看不见的黑矮星。

不是所有的恒星都具有太阳的命运。对于一些恒星，它们的结局会更为激烈。它们的命运为所谓的昌德拉塞卡极限所决定：这就是白矮星所能具有的最大的质量值。根据1929年萨拉玛尼安·昌德拉塞卡的计算，如果恒星的质量大于太阳质量的1.5倍的话，白矮星不能存在。（他是一位年轻的印度研究生候选人，在他从印度到英国的航船上作出这个计算的。）这个计算在1930年左右由苏联的列夫·朗道独立地重复过。现在改善了的昌德拉塞卡极限值大约为

$1.4 \quad M_{\odot}$，

这里$M_{\odot}$代表太阳质量，亦即M⊙等于一个太阳质量。

请注意昌德拉塞卡极限比太阳质量大不了多少。而我们知道许多通常的恒星的质量比这个质量大得多。例如，质量为$2M_{\odot}$的恒星的最终命运是什么呢？根据已有的理论，这些恒星又会肿大变成红巨星，正和前面一样，它的白矮星核会慢慢地得到质量。然而，在某一临界阶段，核质量会达到昌德拉塞卡极限，而泡利不相容原理将不足以抵抗巨大引力所引起的压力 [8] 。在这一点前后，核心灾难式地向内坍缩并遭受到温度和压力的巨大增加。发生了激烈的核反应，从核中以中微子的形式释放出极大的能量。这些把恒星正在向内坍缩的外面区域加热上去，紧接着的是一巨大的爆炸。这个恒星就变成了超新星！

仍在坍缩的核发生了什么呢？理论告知我们，它甚至达到比白矮星惊人的密度还要巨大得多的密度。这核可以作为一个中子星而稳定下来（参阅408页），现在是中子简并压力——也即泡利原理应用于中子——支持着它。它的密度是如此之大，以至于乒乓球大小的体积含有的中子星物质和赫米斯小游星（或者是火星的月亮狄莫斯）一样多。这是原子核中的密度（一个中子星像是一个巨大的原子核，半径大约为几十千米，然而以恒星的标准来看极其微小！）但是，现在有了新的极限

（称为蓝道——奥本海默——沃尔科夫极限），它和昌德拉塞卡极限很类似。其现代（修正）值非常粗略地为

$2.5\quad M_{\odot}$，

若质量超过这一数值，中子星就维持不住。

如果原先恒星的质量足够大，甚至超过这一极限，其坍缩的核会发生什么呢？譬如讲，许多已知恒星的质量的范围是从$10M_{\odot}$ 到$100M_{\odot}$。看来不断地把这么多质量抛出，使剩余下的核的质量低于中子星极限是非常不可能的。与此正相反，预料的结果是会产生黑洞。

什么是黑洞？它是空间或者空间——时间的一个区域，在那里引力场变得如此之强大，甚至连光都不能逃逸。我们记得，相对论原理的一个含义是光速为一极限速度：没有物体或信号能超过局部的光速（参阅251页，272页）。所以，如果光都不能从黑洞逃逸，没有任何东西可能逃逸。

读者或许听说过逃逸速度。这是为了从某一大质量物体逃逸的一个物体必须具有的速度。假设该物体是在地球上，则从地球逃逸的速度为每小时40000千米，也就是大约为每小时25000英里。从地面向任何方向抛出的具有超过此速度值的石头都会完会逃离地球（假定我们忽略空气阻力的效应）。如果以低于此值的速度抛出，则它会落回到地球来。（这样，“任何投掷体必须回归”的命题是不对的；只有它的抛出速度低于逃逸速度时才会返回！）对于木星，逃逸速度为每小时22万千米，也就是大约每小时14万英里；对于太阳为每小时220万千米，或大约为每小时140万英里。假定我们现在想象太阳的质量集中于一个只有现在半径的1/4的球体里，则需要的逃逸速度就要加倍；如果太阳还要更加紧密，比如讲半径减小到1/100，则速度要增大到10倍。我们可以想象对于足够大质量的和足够集中的物体其逃逸速度甚至会超过光速！这种事发生时，我们就有了一个黑洞[9]。

我在图7.13中画出了一个物体坍缩而形成黑洞的空间——时间图（我在这里假定，物体在坍缩的过程中近似地维持着球对称。而且我在这里压缩了空间的一个维数）。我也把光锥画出，正如我们在第5章（参阅268页）讨论广义相对论时所知道的，它们表明物体运动或信号

的绝对限制。我们注意到，光锥向中心倾斜，并且越是靠近中心就越倾斜。

存在一个称作施瓦兹席尔德半径的临界距离，在这距离下光锥的极限在图中变成垂直的。在这距离下光（它必须沿着光锥）只能在坍缩物体上徘徊，光所具有的所有的向外的速度只刚好足以抵抗引力的巨大吸引。在施瓦兹席尔德半径处，这些徘徊的光（也即光线的整个历史）在空间——时间中的轨迹构成的三维面称为黑洞的（绝对）事件视界。任何在视界之内的东西都不能逃离或者甚至都不能和外面的世界通信。这可从光锥的倾斜以及所有运动和讯号被限制在这些光锥之内（或之上）的基本事实看到。由几个太阳质量的恒星坍缩形成的黑洞，其视界的半径为几千米。预料在星系中心存在质量大得多的黑洞。我们银河系的中心很可能包含有一个大约100万太阳质量的黑洞，其视界的半径为几百万千米。

坍缩形成黑洞的实际物质将完全在视界之内完结，而且那时它不能和外界通信。我们简要地考虑一下该物体的可能命运。此刻，我们所关心的仅仅是由它的坍缩产生的空间——时间几何——一种具有极其古怪含义的空间——时间几何。

我们想象一个勇敢（或愚勇）的太空人B，他决心旅行到一个大黑洞中去。而他的胆怯（或谨慎）的同伴A安全地留在事件视界之外。让我们假定A以视线尽可能长久地追踪B的行踪。那么A将看到什么呢？从图7.13可以肯定，A永远见不到B在视界之内的历史（亦即B的世界线）的部分，而A将最终见到B在视界之外的历史部分——虽然在B穿越水平的前一瞬间，由A看起来必须等待越来越长的时间。假定B在自己的钟为12点时穿过视界。A实际上永远见不到这一事件。但在钟表读数为11：30、11：45、11：52、11：56、11：58、11：59、$11:59\frac{1}{2}$、$11:59\frac{3}{4}$、$11:59\frac{7}{8}$ 等时刻，A都能连续地看到（从A的观点看，这大约是发生在相同间隔的时间里）。原则上，对于A而言，B总能被看到，并且显得永远在视界上徘徊，而且B的表在接近这致命的12点钟时显得越来越慢，并永远不能达到这一时刻。实际上，B的影像会非常快地变得朦胧以至无法辨别。这是因为，从B刚好在视界外的世界线的小片段来的光，必须发散在A所经历的余下的时间内。换言之，B在A的视线内消失——这一点也适合于原先的整个坍缩物体。A所能看到的全部的确只能是一个“黑洞”！

关于可怜的B的境况又如何呢？他的经验又如何呢？必须首先指出，当B穿过视界之时，他不会有任何异样的感觉。他凝视着他的手表记录12点钟，他看到了一分钟一分钟规则地流逝。11：57、11：58、11：59、12：00、12：01、12：02、12：03。在12：00点附近似乎没有任何古怪的事发生。他可以回头看A，并发现在整个时间里A总被连续地看到。他可以看到A自己的表。对于B来讲，A的表以一种正常的规则的方式运行。除非B已经计算出他应该穿越这视界，否则他无法知道这一事件 [10] 。视界极端阴险。B一旦穿越进去，就再也逃脱不出来。他的局部宇宙最终会在他周围坍缩，他注定很快就要遭受到自己的“大挤压”。

这也许不是纯粹私人的事务。在某种意义上，形成黑洞的坍缩物体的全部物质首先和他分担“同样”的挤压。事实上，如果在黑洞之外的宇宙是空间闭合的，这样外界的物质也会最终卷入到包罗一切的大挤压中去。那么可以预料，这种挤压和B的“个人”挤压相同 [61] 。

尽管B的命运令人不快，但我们认为，一直到这点为止，他所经历到的局部物理学不会和我们已知并理解的物理学有任何抵触之处。尤其是，我们预料他不会觉得热力学第二定律被违反，更不用说完全反演的熵增加行为了。第二定律在黑洞之内，正如同在其他地方一样成立。B附近的熵仍然增加，直到他的最后挤压为止。

为了理解在（“个人的”或“包罗万象的”）“大挤压”处的熵值确实极高，而在大爆炸处熵低得多，我们还需要进一步研究黑洞的空间——时间几何。但在此之前，读者也应先浏览一下图7.14，该图画出了称作白洞的黑洞的假想时间反演。自然界中也许不存在白洞。但是，它们的理论可能性对于我们具有相当重要的意义。

空间——时间奇点的结构

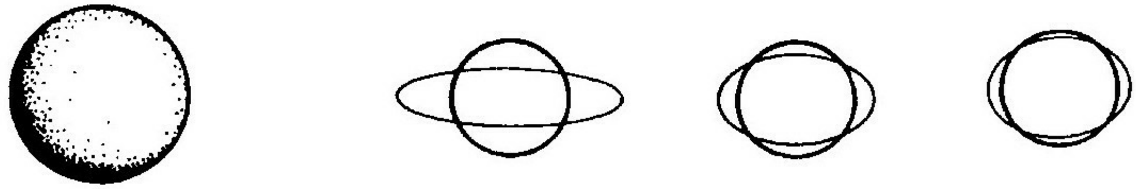


图7.15　随着物体靠近一个球形的引力物体，其潮汐效应按照与物体中心距离成立方反比律的关系增强

我们从第5章263页知道，空间——时间曲率如何以潮汐效应的方式呈现出来。一些在某大物体引力场中自由下落的粒子组成的一个球面在一个方向上被拉伸（沿着朝向引力物体的直线）和在与这垂直的方向上被压挤。这种潮汐效应随着和引力物体的接近而增大（图7.15），其强度变化和离开此物体距离的立方成反比。太空人B在向黑洞内部下落之时便会感到这一增强的潮汐效应。对于一个几倍太阳质量的黑洞，潮汐效应是巨大的——以至于该太空人在靠近黑洞时根本就不能存活，更不用说他穿越黑洞视界了。对于更大的黑洞，视界处的潮汐效应实际上更小。许多天文学家都相信，在银河系中心可能存在一个大约1000000太阳质量的黑洞。当太空人穿越这黑洞视界时，其潮汐效应应该是相当小，虽然也许足以使他稍微感到不舒服。然而，只在太空人穿越之后的很短时间里，这种潮汐效应才继续维持很小。事实上，只要几秒钟的时间它就迅速地达到无限大！不仅这位可怜的太空人的身体会被这一增强的潮汐效应撕开，而且组成他的分子所包含的原子、原子的核，以及最后甚至所有亚原子粒子很快地也难逃厄运！“压榨”正以如此方式施展其终极的淫威。

不仅是所有物体以这种方式被毁灭，甚至空间——时间本身都面临着它的终点！这种最终灾难称作空间——时间奇点。读者一定会问，我们何以知道这种灾难一定会发生，在何种情况下物体和空间——时间注定要遭此厄运。在任何形成黑洞的情形下，这些是从广义相对论的经典方程引出的结论。奥本海默和斯尼德（1939）原先的黑洞模型呈现了这种行为。然而，许多年来天体物理学家总是抱着幻想，认为奇性行为只是在该模型中假定的特殊对称性的孽障。在现实（非对称）的情况下，坍缩的物体也许会以某种复杂的形式旋开并重新逃到外头去，但是，在进行了更一般的数学论证后，这种希望就破灭了。这些论证的结果被称作奇点定理（参阅*Penrose 1965*，*Hawking and Penrose 1970* ）。这些定理断言，在具有合理物质源的广义相对论的经典理论中，引力坍缩情形中的空间——时间奇性是不可避免的。

利用时间方向的反演，我们又类似地发现相应的初始的空间——时间奇性的不可避免性。奇点在任何（适当的）膨胀宇宙中代表大爆炸。此处奇点不代表所有物质和空间——时间最终的毁灭，它代表空间——时间以及物质的创生。在这两种奇点之间也许存在一个准确的时间对

称：初始奇点，空间——时间以及物质在该处创生；终极奇点，空间——时间以及物质在该处消灭。在两者之间的确存在着一个重要的相似，但是在我们仔细考察之后，就会发现它们并非准确的时间反演。它们的几何差异对于我们的理解意义重大，因为它们包含热力学第二定律起源的关键！

让我们回到自我牺牲的太空人B的经验上来。他遭遇到了很快就要增强到无限大的潮汐力。由于他是在空虚的空间中旅行，所以经历了体积守恒和畸变效应，后者是我早先表达成外尔（见第5章264页，270页）的空间——时间曲率张量所提供的。空间——时间曲率张量中代表整体的压缩，并称作里奇的余下的部分在空虚空间中为零。也可能在某一阶段，B在事实上遭遇到物质，但是甚至在发生这种情形时（毕竟他自身是由物质所构成的），我们仍然普遍地发现外尔的测度比里奇大得多。我们预料，接近于最终奇点时的曲率完全是由外尔张量所主宰。一般地讲，此张量趋近于无穷大：

$$外尔\rightarrow\infty,$$

（虽然它会以振荡的形态出现）。这是空间——时间奇点[11]的一般情形。这种行为和高熵的奇点有关联。

然而，大爆炸处的情况与此完全不同。我们早先考虑过的高度对称的弗里德曼——罗伯逊——瓦尔克空间——时间提供了大爆炸的标准模型。外尔张量提供的畸变的潮汐效应在这里完全不存在。取而代之的是作用在检验粒子的球面上的对称的向内的加速度（图5.26）。这是里奇张量而不是外尔张量的效应。在任何FRW模型中，张量方程：

$$外尔=0$$

总是对的。当我们越来越接近这一初始奇点时，我们发现是里奇而不是外尔变成无穷大。这就为我们提供了低熵的奇点。

如果我们在一个准确的坍缩的FRW模型中考察大挤压奇点，我们就会发现在挤压处，外尔=0，而里奇趋于无穷大。然而，这是一种非常特殊的情形，我们不会在完全现实的模型中预料到这种现象。在现实模型

中必须计入引力成团的效应。随着时间的演化，原先以弥散气体的形式存在的物质将结团成恒星的星系。大量恒星将会由于引力收缩而渐序变成：白矮星、中子星和黑洞，以及在星系的中心可能会有的某些巨大的黑洞。这种成团——尤其是在黑洞的情形下——代表了熵的巨大增加（图7.16）。这初看起来使人困惑不解，成团的态代表高熵，而均匀的态代表低熵。我们记得，在一盒气体的情形，成团（譬如所有气体都处于盒子的一个角落）的态具有低熵，而热平衡均匀的态的熵更高。但是考虑到引力，则这一切都被颠倒过来，这是由于引力场的普遍的吸引性质引起的。随着时间的推移，成团现象变得越来越极端，最终凝聚成许多黑洞。它们的奇点联合成极其复杂的终极的大挤压奇点。终极奇点绝不像坍缩的FRW模型中受外尔=0限制的理想大挤压。在所有的时间里，随着越来越多的结团发生，存在外尔张量变得越来越大的倾向[12]。一般来讲，在所有的终极奇点处外尔→∞。图7.17画出了一个代表闭合宇宙和一般描述相一致的整个历史的空间——时间图。

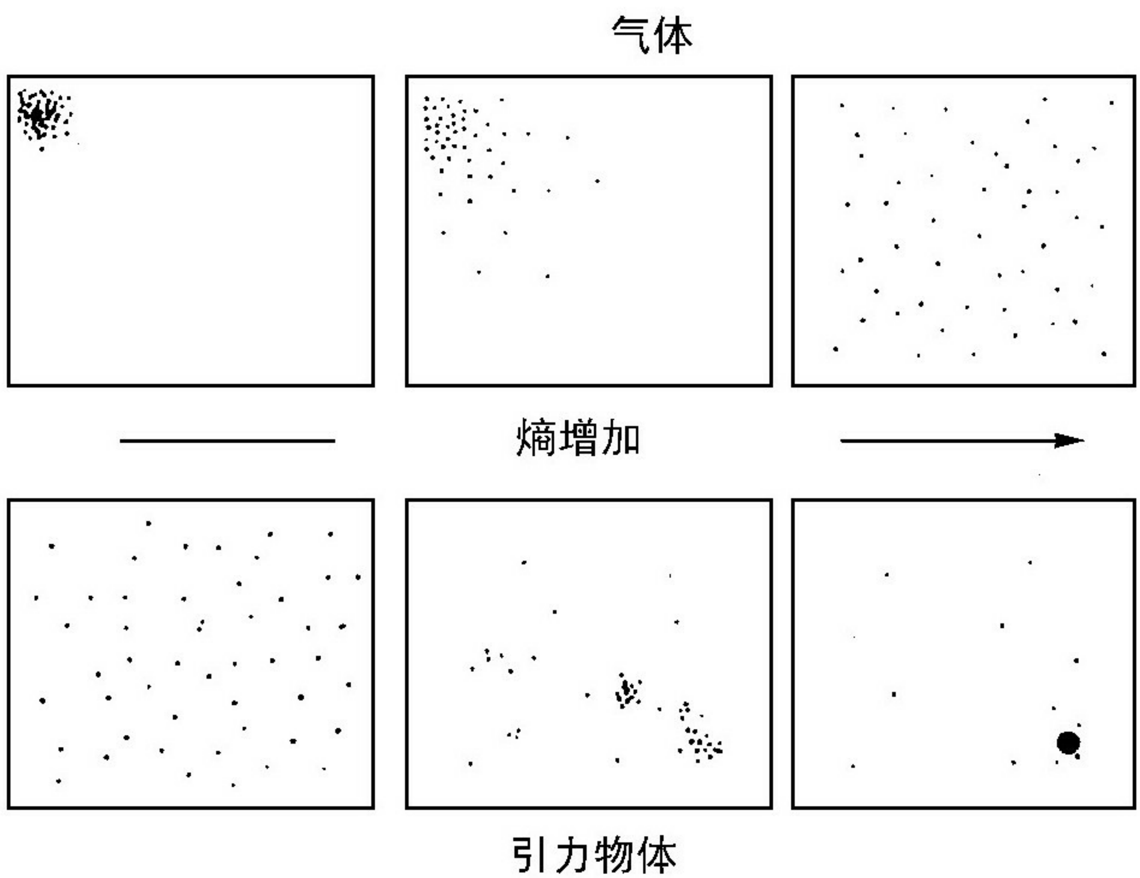


图7.16　对于通常气体，熵增加倾向于使分布更均匀。对于引力物体的系统却是相反的。引力结团引起高熵——最极端的情形是坍缩成一个黑洞

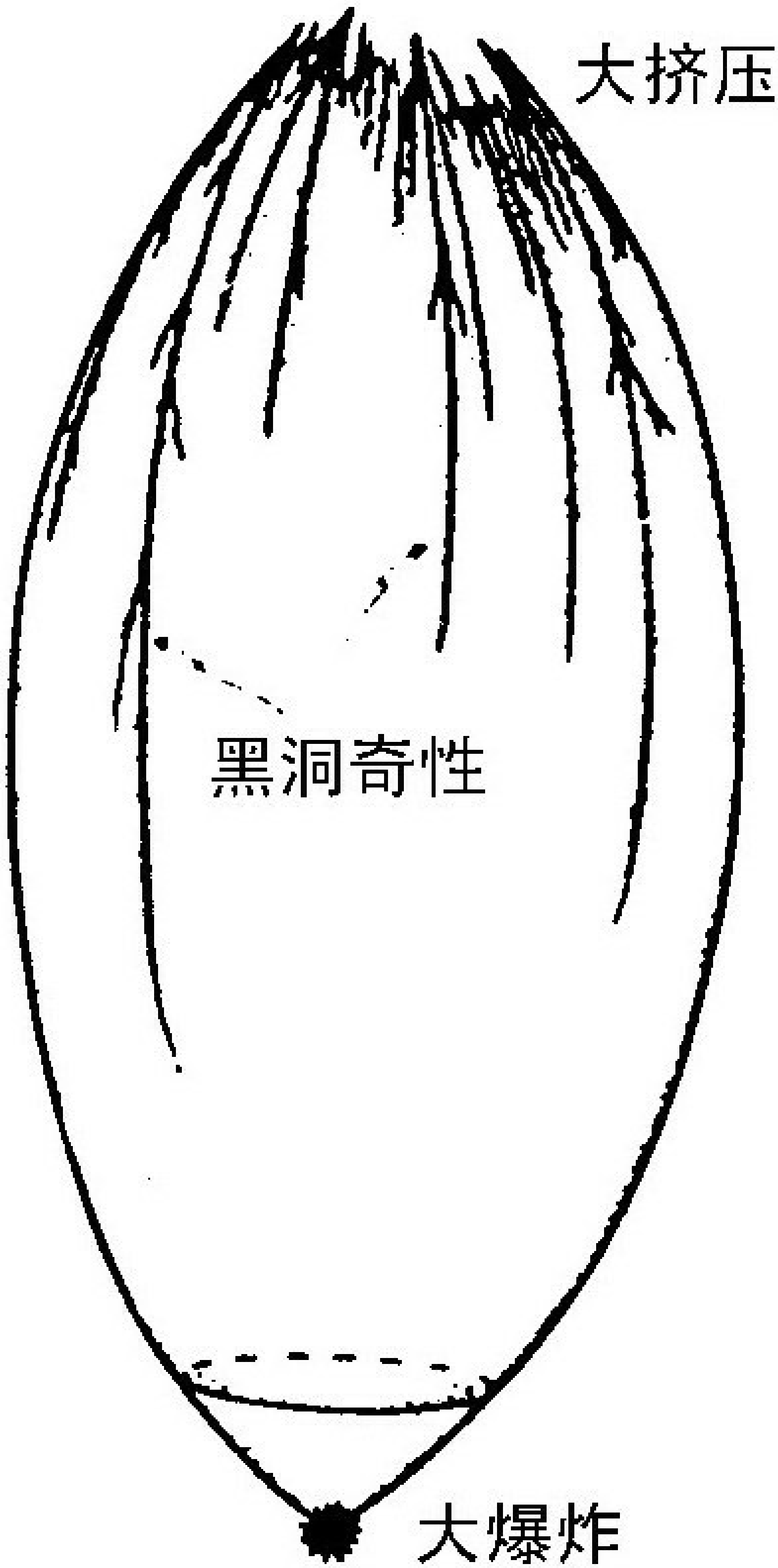


图7.17　一个闭合宇宙的整个历史。这一宇宙是从均匀的低熵的外尔=0的大爆炸开始，终结于

一个高熵的大挤压。大挤压代表许多黑洞的凝聚，并且此时外尔→∞

现在，我们看到一个坍缩的宇宙为何不必具有低熵。大爆炸处的熵的“最低值”为我们提供了第二定律——因此，这不仅仅是大爆炸时刻宇宙“小尺度”的推论！如果我们把上面得到的大挤压图像作时间反演，我们应得到一个具有极其巨大的熵的“大爆炸”，因而不存在第二定律！由于某种原因，宇宙在一种非常特殊（低熵）的态下创生出来，加上的限制有点像在FRW模型中的外尔=0。如果没有这种性质的限制，“更可能”的情况是，初始和终结奇性都具有高熵的外尔→∞的类型（图7.18）。在这种“可能”宇宙中的确不会有热力学第二定律！

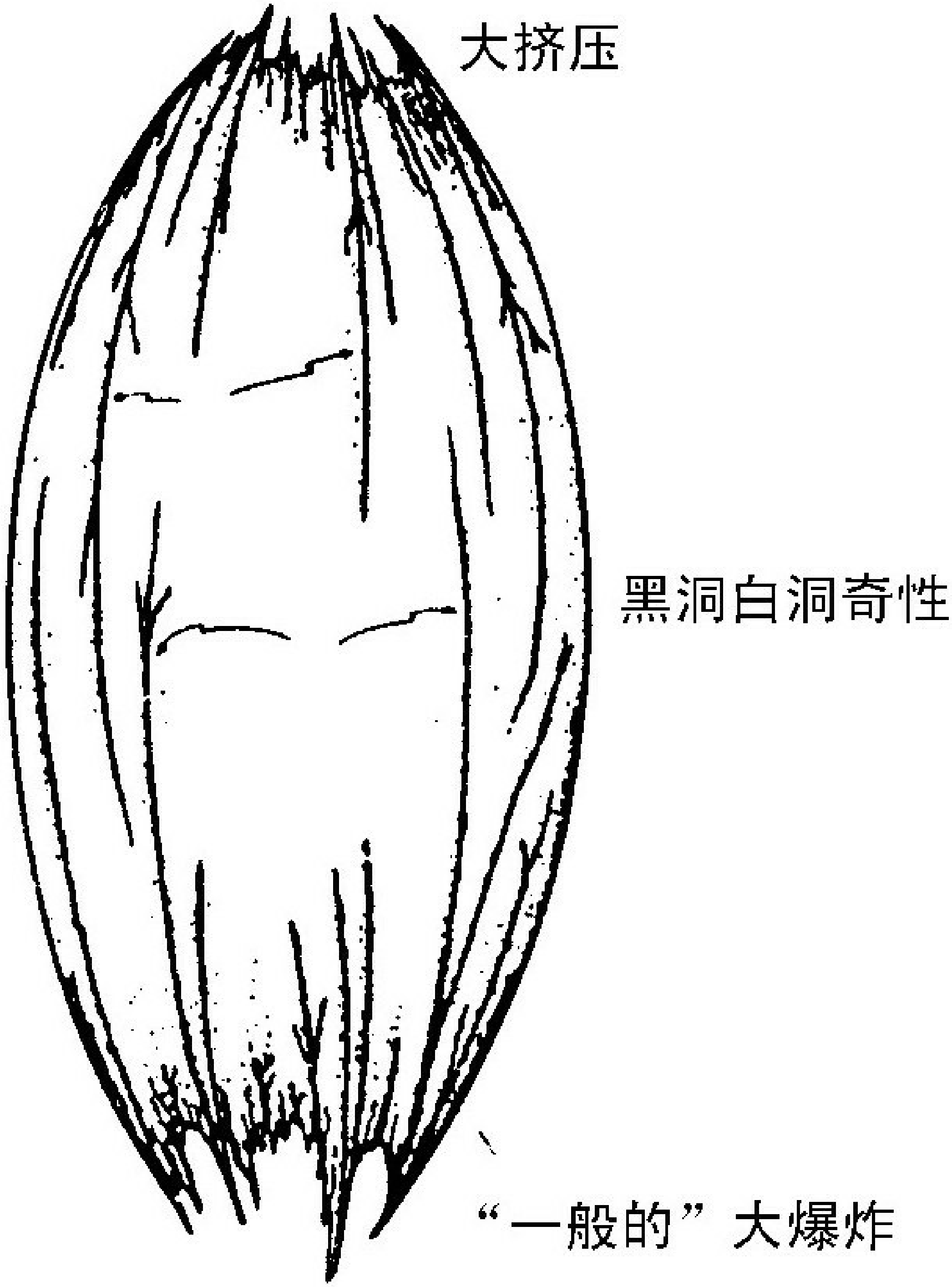


图7.18　如果除去外尔=0的限制，则我们也有高熵的大爆炸，在这里外尔→∞。这样的宇宙会布满白洞，并不存在热力学第二定律，这一切都和常识严重冲突

大爆炸是何等特殊

让我们试图理解在大爆炸处外尔=0的条件下所受到的限制程度。为简单起见（正如上述的讨论），我们假定宇宙是闭合的。为了能得出某些清晰的数字，我们进一步假定在宇宙中质子和中子的总数目，也就是重子数B为：

$$B=10^{80}.$$

除了在观测上B必须至少有这么多以外，（并没有什么选取这一数目的特别理由；有一次爱丁顿断言，他准确地计算出B，其数值和上面的值很接近！似乎再也没有人相信这一特殊的计算，但是这一数值就一直停在10^{80} 上。）如果B的数值取得比这更大（或许实际上B=∞，），那么我们就会得到比现在即将得到的异乎寻常之数字更为惊人的结果！

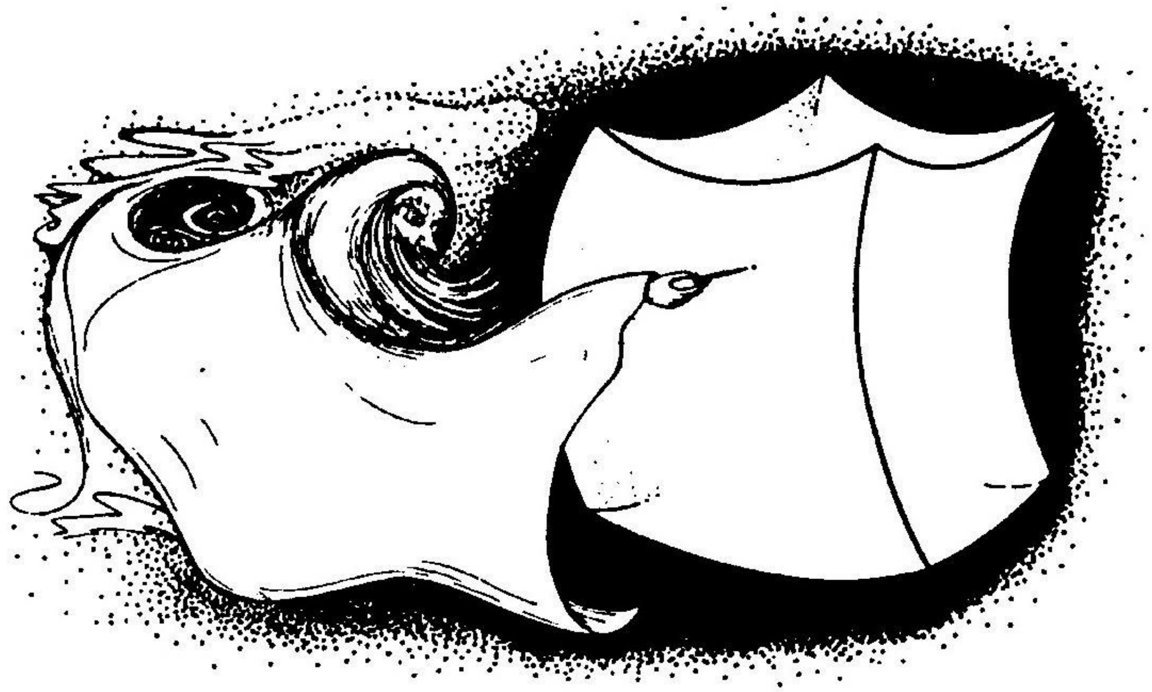


图7.19　为了产生一个和我们生活其中的相类似的宇宙，造物主必须瞄准可能宇宙的相空间中的不可思议的小体积——在所考虑的情形下大约为总体积的

$$1/10^{10^{123}},$$

。（针尖和所瞄准的点不是按比例画出的！）

想象一下整个宇宙的相空间（参阅228页）！这一相空间中的每一点代表宇宙启始不同的可能方式。我们可以想象造物主，它把一个针尖

点在相空间中的某一点上（图7.19）。针尖的不同位置提供不同的宇宙。而造物主目标所需的精度决定于它所创造的宇宙的熵。由于相空间的巨大体积可让针尖去戳，所以产生一个高熵宇宙是相对“容易”一些。（我们记得熵和有关相空间的体积的对数成正比。）但是，为了使宇宙从低熵态起始——以保证存在热力学第二定律——造物主必须瞄准相空间中极其细微的体积。为了使结果和我们生活其中的宇宙相类似，这一区域应该是多小呢？要回答这个问题，首先必须先提到一个非常出色的公式，由雅科布·柏肯斯坦（1972）和史蒂芬·霍金（1975）所发现计算黑洞的熵的公式。

考虑一个黑洞，并且假定其视界面积为A。柏肯斯坦——霍金黑洞熵公式则为：

$$S_{\mathrm{bh}} = \frac{A}{4}\left(\frac{kc^3}{G\hbar}\right),$$

此处k为玻耳兹曼常数，c为光速，G为牛顿引力常数，$\hbar$ 为普朗克常数除以2π。此公式的主要部分为A/4，括号内的部分只不过是包括了合适的物理常数。这样，黑洞的熵和它的表面积成正比。对于一个球对称的黑洞，此表面积和黑洞质量的平方成正比：

$$A = m^2 \times 8\pi\ (G^2/c^4)。$$

把它和柏肯斯坦——霍金公式合并，我们就看到黑洞的熵和它的质量平方成比例：

$$S_{bh}=m^2\times 2\pi（kG/\hbar c）。$$

这样，黑洞的单位质量的熵（S_{bh} /m）和它的质量成正比，所以黑洞越大，它就越大。因此，对于给定的质量，或由于爱因斯坦的公式 $E=mc^2$ 而等效的能量，当物质坍缩成一个黑洞时获得最大的熵！而且，当两个黑洞相互并吞而产生一个单独黑洞时得到巨大的熵！诸如在星系中心发现的那些巨大的黑洞能提供极其了不起的熵值——远比在其他类型的物理情形下遇到的熵值大得多。

只需要很少的条件就可断言，当所有质量都集中到一个黑洞中时得到的熵最大。霍金的黑洞热力学分析指出，必须有一非零的温度和黑洞相关联。其中的一个含义便是，并非所有的质量能量都包含在黑洞之中。在最大熵状态下，最大熵是在一个黑洞和“辐射的热库”相平衡时才获得。对任何合理尺度的黑洞，这辐射温度实在非常小。例如，一个太阳质量的黑洞，其温度大约为10-7开，这比迄今为止在任何实验室里所能测量到的最低温度还要低，比星际空间的2.7开温度低得多了。对于更大的黑洞，其霍金温度甚至还要更低！

只有在下面两种情形下霍金温度对于我们的讨论才有意义：（i）也许在我们宇宙中存在称为微黑洞的微小得多的黑洞；或（ii）宇宙不在霍金蒸发时间，也就是黑洞完全蒸发所需的时间之前坍缩。关于（i），微黑洞只能在适当的混沌大爆炸时产生。实际中这类微黑洞不会大量存在，否则它们的效应应该已被观测到；而且，依我的观点，它们根本就不存在。关于（ii），对于太阳质量的黑洞，霍金蒸发时间大约为目前宇宙年龄的10^{54} 倍。对于更大的黑洞，其时间要长得多。这些效应似乎不会根本改变上述的论断。

为了对黑洞熵的巨大数值有一概念，我们可以考虑原先以为对宇宙熵有最大贡献的2.7开的黑体背景辐射。这种辐射所包含的巨大数量的熵慑服了天体物理学家，它远远超过人们在任何其他过程（例如在太阳中）所遭遇到的通常的熵值。背景的熵大约是每一个重子10^8 （此处我选用“自然单位”，这样玻耳兹曼常数为1），（实际上，这表明每个重子在背景辐射中对应于10^8 个光子。）所以，如果共有10^{80} 个重子，则

我们宇宙中的背景辐射应有总熵：

10^{88} .

如果没有黑洞，这一数值的确代表了宇宙的总熵，由于背景辐射中的熵淹没了所有其他通常过程的熵。例如太阳中的每个重子的熵的数量级为1。另一方面，按照黑洞的标准背景辐射的熵是微不足道的。柏肯斯坦——霍金公式告知我们，在太阳质量的黑洞中每一重子的熵大约在 10^{20} 自然单位左右。这样，要是宇宙全部由太阳质量黑洞所构成，则总数值会比上面给出的大许多，也就是：

10^{100} .

当然，宇宙不是这样构成的，但是这一数值开始告诉我们，当计入引力的无情效应时，背景辐射的熵是如何地“微小”。更现实一点，假定我们的星系不完全由黑洞组成，而主要由通常的恒星组成，在包含10^{11}个通常恒星的星系的核中假如有个1000000（亦即10^{6} ）太阳质量的黑洞（这对我们自己的银河系是合理的）。计算结果指出，现在每个重子的熵实际上比前面巨大的数值还要大，也就是10^{21} 。这样，以自然单位给出的总熵为：

10^{101} .

我们可以期待，在非常长的时间后，星系质量的主要部分会被并吞到它们中心的黑洞中去。发生此事之后，每一重子的熵变为10^{31} ，其总熵具有极大的数值：

10^{111} .

然而，我们是在考虑一个闭合的宇宙，这样它将最终坍缩；所以，似乎整个宇宙形成一个黑洞。可以合理地利用柏肯斯坦——霍金公式估计最终大挤压的熵。这就给出了每一重子10^{43} 的熵，而整个大挤压的无与伦比巨大的总熵为：

10^{123} .

这一个数值给出了造物主所能得到的相空间总体积的估计。熵应该表达成最大区域体积的对数。由于10^{123} 是该体积的对数，所以其体积按自然单位应为10^{123}的指数，也就是：

$$V=10^{10^{123}}$$

（某些聪明的读者会觉得我应该用数值$e^{10^{123}}$，但是对于这么大的数，e和10在本质上是可互相取代的！）为了给我们提供一个和热力学第二定律以及我们现在所观察的相一致的宇宙，造物主必须瞄准的原先的相空间体积W应为多大呢？我们取下面的两个数值中的任一个根本关系不大：

$$W=10^{10^{101}} 或 W=10^{10^{88}},$$

它们分别为星系黑洞或者背景辐射给出的数据，或是在大爆炸处更小得多（事实上更为合适）的实在的数据。不管哪种数值V和W的比率接近于：

$$V/W=10^{10^{123}}。$$

（试试看，$10^{10^{123}}\div 10^{10^{101}}=10^{10^{123}-10^{101}}=10^{10^{123}}$，非常接近。）

这就告诉我们造物主要瞄得多准：也就是要准确到$1/10^{10^{123}}$。这是一个异乎寻常的数值。人们甚至不能把这个数以通常十进位的办法完全写下来。它是1后面连续跟10^{123} 个0！甚至如果把0写在整个宇宙中每一颗单独的中子和质子上——还可以加上所有其他的粒子——人们发觉还是远远不够写下所需要的这一个数值。使宇宙准确地运作所需要的精度，比制约从一个时刻到另一时刻事物行为的任何超等动力方程（牛顿、麦克斯韦、爱因斯坦的方程）我们已习惯的精度毫不逊色。

但是，为何大爆炸是如此精密地策划的，而大挤压（或黑洞中的奇

点）却是预料中完全混沌的呢？这可按照在空间——时间奇点处的空间——时间曲率的外尔部分的行为来重述这个问题。我们发现在初始的而不是终结的奇点处存在约束：

外尔=0

（或某种和它非常类似的东西）。似乎正是这个限制造物主选择相空间内这个非常微小的区域。这限制适合于任何原始（而非终结）的空间——时间奇点。我把它称作外尔曲率假设。这样，如果我们要理解第二定律从何而来，似乎就必须理解为何这样的一个时间反对称的假设必须成立[13]。

我们如何才能对第二定律的起因有更深入的理解呢？我们似乎被逼迫到死路上去。我们必须理解为何空间——时间奇点具有它所具有的结构；但是空间——时间奇点是我们物理理解达到极限的区域。有时人们把空间——时间奇点存在所导致的死胡同和另一事件相提并论：那就是20世纪初物理学家研究原子稳定性（参阅289页）所遭遇到的困难。在每种情况下，早已确立的经典理论总是得出“无穷大的”答案，因而经典理论对于这样的使命无能为力。量子理论阻止了原子电磁坍缩的奇异行为，正如量子理论应在恒星的引力坍缩“无限的”经典空间——时间奇点处得到有限的理论。但是这绝不是通常的量子理论。它必须是空间和时间结构本身的量子力学。这样的理论，若存在的话，应称为“量子引力”。量子引力还不存在并非因为物理学家不努力，或者没有专长和天才。许多第一流的科学头脑专心致志于建立这样的理论，可惜未成功。这是我们试图理解时间流逝的方向性时所最后面临的绝境。

读者一定会问，我们经历了什么样的旅途。在我们追求理解为何时间显得只向一个方向而不向另一方向流逝的过程中，我们已经旅行到时间的最终点，在该处空间概念本身都被瓦解了。我们从这一切得到了什么教益呢？我们发现理论还不足够于提供答案。但是，这对我们试图理解精神又有什么用场呢？尽管缺少足够的理论，我相信我们的航程的确给予我们重要的教导。现在我们应该回过头来。我们的归程将比出发更加冒险，但是依我看，没有其他合理的归途！

第8章　量子引力的寻求

为什么需要量子引力

我们在前一章所了解的有哪些是和大脑与精神相关的呢？虽然我们瞥见了某些作为我们认知“时间流逝”方向性质、相互纠缠的基础物理原则，我们似乎迄今并未洞察到为何我们感觉时间流逝或甚至我们为何感觉得出时间。依我看来，这里需要激进得多的观念。虽然我有时强调的地方与众不同，迄今我的陈述并不特别激进。我们已经熟悉了热力学第二定律，我已试图说服读者，大自然以她所选择的特殊形式呈现在我们面前的这一个定律，其起因可以追溯到宇宙的大爆炸起源时无与伦比的几何限制：外尔曲率假设。有些宇宙学家宁愿以某种不同的方式来表达这个初始限制，然而这种对初始奇点的限制无论如何是必需的。我准备从这个假设抽取的推论将比此假设本身更激进。我宣称，量子理论的框架本身需要一个变革！

这个变革会对量子力学和广义相对论进行适当的统一，也就是在寻求量子引力时起作用。大多数物理学家相信，量子力学在和广义相对论统一时不需要变革。而且他们争辩道，在与我们大脑相关的尺度下任何量子引力的物理效应必然都是毫无意义的！他们会（非常合情理地）说，虽然这些物理效应在极其微小的尺度下的确重要，这尺度被称为普朗克长度[62]—等于10^{-35}米，比最小的亚原子、粒子的尺度小一万亿亿倍——但在非常大的“寻常”尺度，好比说只有10^{-12}米时，这些效应与所发生的现象没有直接关联，这种尺度下的现象是由对大脑活动很重要的化学或电作用支配着。的确，甚至经典的（也就是非量子的）引力对这些电和化学活动几乎没有影响。如果经典引力都没有效应，为什么对经典理论的微小的“量子修正”会产生任何重大的差异呢？况且，由于量子理论引起的偏差从未被观察到，如下的推断就显得更不合理，即对标准量子理论的任何假想的微小偏差在精神现象中会起任何可以想象得到的作用！

我的论证与此非常不同。我并不这么关心量子力学对空间——时间结构理论（爱因斯坦广义相对论）的效应；而是相反的，也就是爱因斯

坦的空间——时间理论在量子力学结构本身的可能的效应。我应强调指出，我现在所提出的是一种非传统的观点。正是在广义相对论对于量子力学的结构具有影响这一点上是非传统的！传统的物理学家总是非常不情愿去相信，量子力学的标准结构应在任何方面被擅改。虽然把量子理论的规则直接应用于爱因斯坦理论的确遭遇到了似乎不可克服的困难，该领域工作者的反应是企图用它作为修正爱因斯坦理论的理由，而不是修正量子理论 [1] 。我本人的观点几乎是完全相反的。我相信量子理论自身的问题具有基本的特征。我们记得量子力学中两个基本过程U和R之间的非一致性（U服从完全决定性的薛定谔方程——称作幺正演化——而R为随机的态矢量减缩，只要人们认为进行了一次“观测”，则必须经历这样的一个过程）。依我看来，这种不协调性不能仅仅靠采取适当的量子力学“解释”予以解决（虽然普遍的观点似乎认为这样可以）。它只有在某种激进的新理论的框架中才能被解决，而这两种过程U和R被认为是对于包容更广的、更精确的单独过程的不同的（而且非常优越的）近似。所以，我的观点是，甚至这不可思议地精密的量子力学理论都必须改变，而其改变的性质的强烈暗示必须来自于爱因斯坦的广义相对论。我甚至还可进一步断言，在实际上正是所寻求的量子引力理论应该包含这一想象中结合的U/R步骤作为它的基本要素。

另一方面，按照传统的观点，量子引力任何直接的含义都具有更神秘的性质。我提到过，空间——时间结构在不可思议的普朗克长度的极小尺度下会有基本改变。还有人相信（依我看来，这已被确证）量子引力和近年观测到的“基本粒子”的整个家族性质的最终确定在根本上是相关联的。例如，现在不存在任何解释粒子质量为何必须这么大的好理论——而“质量”是和引力概念密切相关的概念。（质量的确是唯一的引力的“源”。）此外，许多人预料（根据1955年瑞典物理学家奥斯卡·克莱因提出的观念）正确的量子引力理论应当可以消除折磨着传统量子场论的无限大（参阅371页）。物理学是一个整体，当我们最终得到真正的量子引力理论时，它肯定应包含我们对大自然普遍定律详细理解的根本部分。

然而，我们距离这样的理解还很遥远。并且所推想的量子引力理论一定和制约大脑行为的现象相距非常远。在解决上一章遭遇到的困难——空间——时间奇点的问题时必需的量子引力（普遍承认的）作用和大脑活动之关联显得特别遥远。这是爱因斯坦经典理论在大爆炸、在黑洞中，以及大挤压所引起的奇点，如果我们的宇宙注定最终要坍缩的

话。是的，这个作用似乎是遥远的。然而，我将论断，这里存在一个无从捉摸但却很重要的逻辑联结网络。让我们考察这个联结是什么样子的。

外尔曲率假设的背后是什么

正如我在前面提及的，甚至是传统的观点告诉我们，必须用量子引力来辅助广义相对论的经典理论来解决空间——时间奇点之谜。这样，量子引力就在经典理论得出没有意义的“无穷大”的答案之处，为我们提供了某种条理一贯的物理。我肯定同意这种观点：这的确应是量子引力留下标志的显明之处。然而，理论家们似乎还没有充分地接受如下惊人的事实，即量子引力的标志公然是时间不对称！在大爆炸——也即过去的奇点处——量子引力应该告诉我们，一个像

外尔=0

的条件必须在按照经典的空间——时间几何概念来描述成为有意义的时刻成立。另一方面，在黑洞内部的奇点以及在（可能的）大挤压——未来的奇点——处却并没有这样的限制。我们预料，当靠近这种奇点时外尔张量变成无穷大：

外尔→∞。

依我看来，这非常清楚地表明，我们所寻求的实际的理论应该是时间不对称的！

我们所寻求的量子引力必须为一个时间不对称的理论。

我必须警告读者，尽管从我所陈述的方式来看，这一结论显然是必需的，却没被当作智慧被接受！许多在这领域的工作者对此采取迟疑的态度。其原因似乎在于，没有一种清晰的方法使得传统的，被充分理解的（就目前进行的）量子化步骤能产生一个时间不对称[2]的量子理论，而这些步骤所应用的经典理论（标准的广义相对论或它的某种流行的修正形式）本身是时间对称的。相应地，这样的量子引力家在考虑这些问题时——这是罕见的！）就需要往他处寻求大爆炸处的低熵的“解释”。

也许，许多物理学家会争论道，一个像初始外尔曲率的零值的假设是被当作“边界条件”的选取而不是动力学定律，它并不在物理学所能解释的能力之内。他们在实际上是论断，一次“上帝的行动”把边界条件赋予了我们，我们不能企图去理解何以我们被赋予这一种而非那一种边界条件的问题。然而，正如我们已经看到的，这个加在“造物主针尖”的限制条件，其非凡与精确绝不亚于我们现在了解的牛顿、麦克斯韦、爱因斯坦、薛定谔、狄拉克及其他精密而优雅的动力学方程。虽然热力学第二定律似乎具有模糊和统计的特征，但是它是由具有无与伦比的精密的几何限制所产生的。有一种观点认为，人们无望理解作用于大爆炸处的“边界条件”的限制。而科学手段却在理解动力学方程上显得如此有价值。这对我来说似乎是不可理喻的。依照我的思维方式，虽然前者是科学迄今不能适当理解的部分，它正和后者一样同为科学的一部分。

科学史已为我们显示出，把物理的动力学方程（牛顿定律、麦克斯韦方程等）和这些所谓的边界条件——也即为了从这些方程的大部分不适合的解中挑出适合的一个解而附加的条件——分开是多么有价值的思想。动力学方程在历史上找到了简单的形式。粒子运动满足简单的定律，但是在宇宙中和我们共存的粒子的实际形态通常不很简单。有时这种形态初看起来简单——诸如行星运动的椭圆轨道，正如开普勒所肯定的那样——但是后来发现。它们的简单性是动力学定律的推论。更深刻的理解总是通过动力学定律才会得到，而如此简单的形态总是更复杂的形态的近似，譬如实际观测到的受扰动的（不完全椭圆的）行星运动。牛顿动力学方程对所有这些都能予以解释。初始条件用以“启动”问题中的系统，而动力学方程从那一时刻开始接手。我们能把动力学行为和宇宙实际内容的形态问题分开是物理科学一个最重要的成就。

我讲过，这种把动力学方程和边界条件的相分离，在历史上具有极大的重要性。进行这种分离的可能性，似乎总是由物理学中引起的特殊类型方程（微分方程）性质的结果。但是我不相信这种分离总是成立。我相信，当我们最后理解实际上制约我们宇宙行为的定律或原则，而不仅仅是我们逐步理解的不可思议的近似，也即构成迄今为止的超等理论之时——我们就会发现，这种在动力学方程和边界条件之间的差别将消失殆尽，而代之以仅仅是某种无比美妙的、协调的、广泛的方案。我在讲到这些时当然只是表达非常个人的观点。其他人也许不同意。但是，这正是在我探讨某种量子引力的未知理论的含义时，在我脑袋中模糊地出现的观点。（这个观点还将影响最后一章一些更富于猜测性的思

考。）

我们怎样才能探讨一个未知理论的含义呢？事情也许并不如它们初看起来那么毫无希望。关键在于一致性！首先，我要求读者接受我们想象的理论——我把它称之为CQG“正确的量子引力”！—会给外尔曲率假设（WCH）提供一个解释。这表明初始的奇点必须在它的立即的未来受外尔=0的限制。这个限制应为CQG定律的推论。它必须适用于任何“初始奇点”，而不仅仅适合于我们称之为“大爆炸”的特殊奇点。我并不是讲，在我们的实际宇宙中除了大爆炸外需要有任何其他奇点。其关键在于，如果还有，则任何这样的奇点必须受到WCH限制。原则上来说，一个初始奇点是粒子可以从那里出来的地方。这和黑洞奇点的行为刚好相反。黑洞奇点是终极奇点——粒子可能落到里面去。

一种可能和大爆炸不同的初始奇点类型是白洞里的奇点。正如我们在第7章讲到的，白洞是黑洞的时间反演。但是我们知道黑洞里的奇点满足外尔→∞。这样对于白洞，我们也必须有外尔→∞。但是，现在的奇点为一初始奇点，对于初始奇点WCH要求外尔=0，这样WCH排除了在我们宇宙中白洞发生的可能性！（幸运的是，这不仅仅是基于热力学的要求——因为白洞会严重地违背热力学第二定律——它也和观察不一致！每隔一阵，总有不同的天体物理学家假想白洞的存在用以解释某个现象，但是这样做总是引起比要解决的问题更多的问题。）请注意，我不把大爆炸本身称作“白洞”。一个白洞具有定域的，不满足外尔=0的初始奇点。但是包容一切的大爆炸，假定它的确被WCH限制，能够满足外尔=0，因而是允许存在的。

还存在另一种“初始奇性”的可能性：亦即黑洞爆炸的那一点。譬如讲，黑洞在10^{64} 年长的霍金蒸发后消失（参阅436页，还有后面的458页）！关于这个假想（似乎被论证得头头是道的）现象的准确性质有许多猜测。我想，这似乎和WCH不矛盾。这样的一个（定域的）爆发实际上可以是瞬息地对称的，我认为和外尔=0的假设没有冲突。无论如何，如果不存在微黑洞（参阅436页），很可能是直到宇宙存在了比现在年龄T长10^{54} 倍以后才第一次发生这类爆发！为了估计$10^{54}\times T$究竟多长，想象把T压缩到能被测量的最短的时间——任何不稳定粒子的最微小的衰变时间——则我们现存宇宙的年龄还比在这个尺度上的$10^{54}\times T$小1万亿倍以上！

有些人采用和我不同的看法。他们论证道[3]，CQG不应为时间不对称的。它在实际上允许两类奇点结构，一种需要外尔=0，而另一类允许外尔→∞。我们宇宙中的刚好是第一类奇点，而我们对时间方向的感觉（由于继而引起的第二定律），把这个奇点放置在我们称之为“过去”而不是“将来”之处。然而，我觉得就这样论证是充分的。它没有解释为何既没有别的外尔→∞类型（也没有别的外尔=0类型）的初始奇点。依照这种观点，为什么宇宙中并没有缀满着白洞？由于宇宙被假定缀满了黑洞，我们需要解释为何不存在白洞[63]？

关于这一点，人们有时祈求所谓的人存原理（参阅*Barrow and Tipler 1986*）。根据这种论证，我们所观察到的，所居住的特殊宇宙是从所有可能的宇宙中由以下事实挑选出来的，这就是我们（或至少某种有知觉的动物）需要存在那里实际地对其观察！（我将在第10章再讨论人存原理。）利用这种论证，人们断言，智力生命只能居住在非常特别类型的大爆炸宇宙中，所以诸如WCH应为这个原则的推论。然而，这种论证不可能得到接近于大爆炸的“特殊性”所需的，在第7章得到的数值$10^{10^{123}}$（参阅437页）。通过非常粗略的计算得出，整个太阳系和它所有的居住者可简单地用粒子随机碰撞而更“便宜”得多地产生，也就是说，其“不可能性”（以相空间体积来测量）比$1/10^{10^{60}}$大得多。这就是人存原理能为我们所做的一切。我们仍然短缺所需要的数值。况且，正如前面刚讨论的观点，人存原理不能为不存在白洞提供解释。

态矢量减缩的时间不对称

我们似乎的确得到结论，CQG必须是一种时间不对称的理论，而WCH（或某种类似物）为这个理论的一个推论。从两个时间对称的部分：量子理论和广义相对论，怎么得到一个时间不对称的理论呢？存在一些可达此目的合情合理的技术可能性，但没有一种可能性被充分地探索过（参阅*Ashtekar et al.1989*）。然而，我希望考察另一途径。我曾经指出，量子理论是“时间对称的”，但是这实在是只适合于该理论的U部分（薛定谔方程等）。我在第7章开头讨论物理定律的时间对称性时，故意不理会R部分（波函数坍缩）。似乎有一种流行的观点认为，R部分也应为时间对称的。产生这种观点的部分原因也许是在把R当作和U相独立的实际步骤这一点上迟疑不决，这样子U的时间对称应意味着R也具有时间对称。我想论断道这是不对的：R是时间不对称的——至少

在我们如果完全把R当作物理学家在计算量子力学的概率时所采取的步骤时是这样的。

首先让我提醒读者应用于量子力学中的称为态矢量减缩（R）的步骤（回顾图6.23）。我在图8.1中示意地画出了态矢量演化的奇怪方式。大部分时候，其演化是依照幺正演化U（薛定谔方程）。但在不同的时刻，当认定进行了“观察”（或“测量”）时，就要采取步骤R，这时态矢量|Ψ〉跃迁到另一个态矢量，例如|x〉，这儿的|x〉是所进行的特别观测O的性质决定的、正交的、两个或更多个不同的可能性|x〉，|φ〉，|θ〉……中的一个。现在，从|Ψ〉跃迁到|x〉的概率由|Ψ〉长度平方|Ψ〉^{2} 在|Ψ〉的希尔伯特空间的|x〉方向投影时减少了的量决定。（在数学上，它和|〉在|Ψ〉方向投影时$|x|^2$ 所减小了的量一样。）这一步骤是时间不对称的。因为紧接着进行观察O以后，由O所决定的不同选择如|x〉，|φ〉，|θ〉……给定集合，态矢量为其中之一，而在O以前的那一时刻，态矢量为|Ψ〉，它不必为这些给定的选择之一。然而，这一非对称只是表面的，它可由对态矢量演化采用不同的观点而得到补救。让我们考虑量子力学的时间反演的演化。这个古怪的描述可用图8.2来说明。现在我们的态在O之前而不是之后的瞬息为|x〉，我们在时间往回的方向上直到前一观察O'那个时刻止应用幺正演化。我们假定往回演化的态变成|'〉（在紧拉着观察O'之后的将来）。在正常向前演化的图8.1的描述中，刚在O'的未来我们具有其他的某个态|Ψ'〉（在正常描述中，观察O'的结果|Ψ'〉向前在O处演化成|Ψ〉）。现在，在我们反演的描述中，态|Ψ'〉还起一个作用：它代表在O'之前的那一时刻系统的态。态矢量|Ψ'〉是实际在O'处观测到的态，这样按照我们反演化的观点，|Ψ'〉变成在时间反演意义上在O'处观测到的“结果”的态。联结O'处观察和O处观察的结果的量子概率P'的计算由态Ix'〉在方向|Ψ'〉上投影时$|x'|^2$ 减小的量给出（这和|Ψ'〉投影到|x'〉上时$|\Psi|^2$ 的减小是一样的）。事实上，这正是我们以前得到的同一个数值，它是U运算的基本性质[4]。

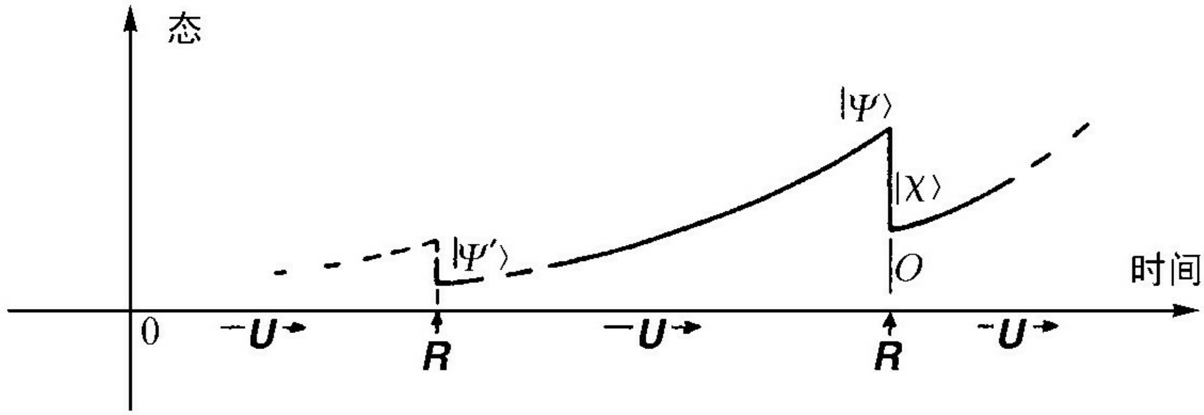


图8.1　态矢量的时间演化：光滑的幺正演化U（服从薛定谔方程）为不连续的态矢量减缩R所打断

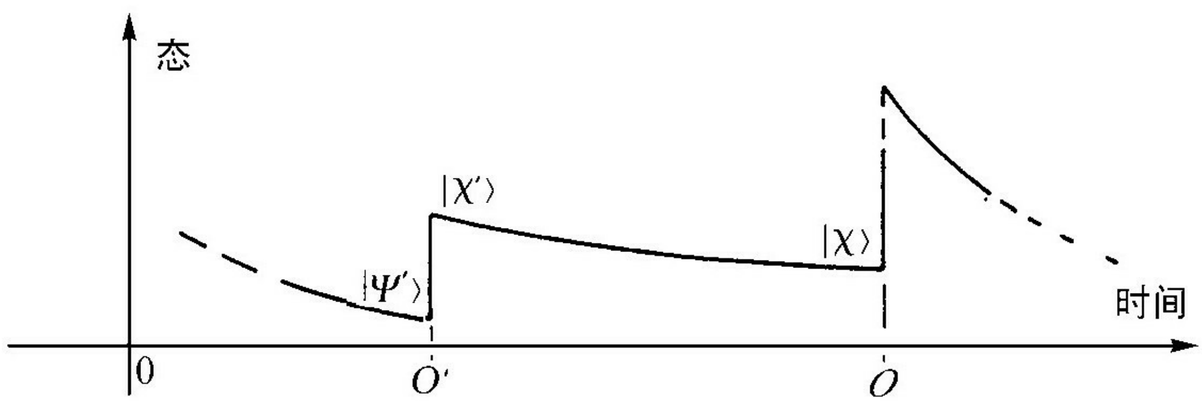


图8.2　态矢量演化的更怪异的图像。此处使用时间反演的描述，联结在O处和O'处观测所计算的概率和图8.1中一样，但该计算值的含义又是什么呢？

这样，甚至在考虑除了通常的幺正演化U以外的由态矢量减缩R描述的不连续过程后，我们似乎确认了量子理论是时间对称的。然而，情况并非如此。用任一种方式计算，量子概率P所描述的是当给定O'处的结果（亦即|Ψ'〉 ）时在O处找到结果（亦即|x〉 ）的概率。这不必与当O处给定结果时在O'找到结果的概率相同。后者 [5] 正是我们时间反演的量子力学应该得到的。令人奇怪的是，这么多的物理学家都暗中假定这两个概率是一样的。（我自己也因为这个假定感到内疚，参阅*Penrose 1979b* ，P584。）然而，这两个概率很可能极其不同。事实上，量子力学只是正确地给出了前者！

让我们在非常简单的特殊情况下考察这一问题。假设我们有一个灯泡L和一个光电管（亦即光子探测器）P。在L和P之间安装有一面半镀

银镜子M，它对LP连线倾斜某一角度，譬如45°（图8.3）。假定灯泡以某种随机的方式不时偶尔发射出光子。因为灯泡的构造（人们可用抛物反射镜）使得这些光子总是非常精确地瞄准着P。只要光电管接受到一个光子就记录下这个事件。我们还假定有百分之百的可靠性。还可假设，只要光子发射出，这个事实就在L也以百分之百的可靠性被记录下来。（在这些理想的要求中，没有任何和量子力学原则相冲突之处，虽然使这些达到这等效率也许是困难的。）

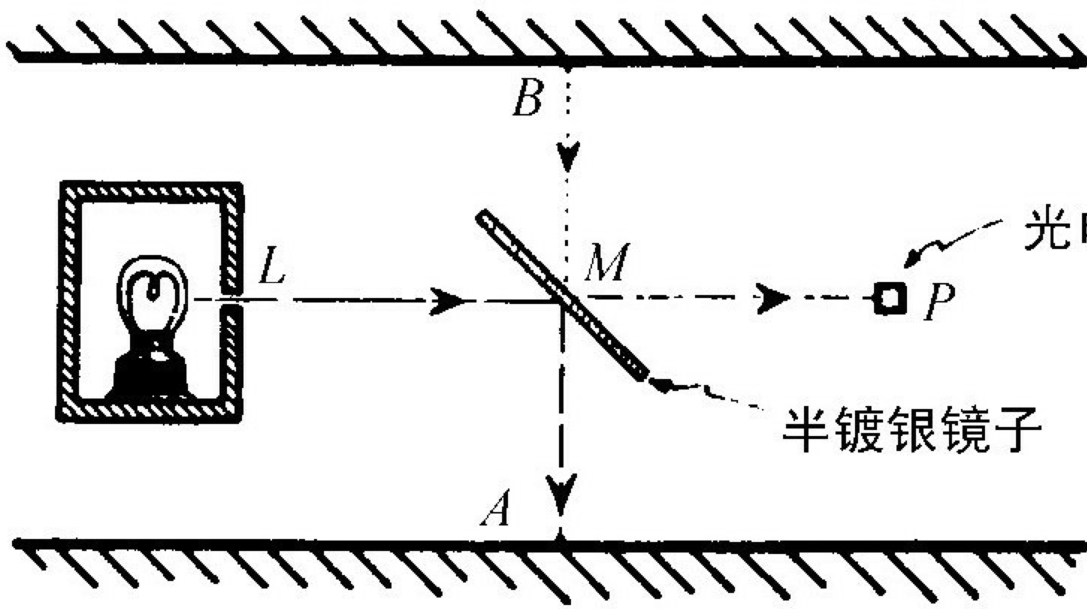


图8.3　一个简单的量子实验中，R的时间不可反演。在光源发射一个光子时，光电管检测到一个光子的概率刚好是一半；但是假定光电管检测到一个光子时，光源发射出一个光子的概率肯定不会是一半

半镀银的镜子M刚好把打到它上面的光子的一半反射，并让另一半穿透过去。更准确地讲，我们必须按照量子力学来思考。光子波函数发射到镜子上，并被分成两半。反射波的幅度为$1/\sqrt{2}$，而透射波的幅度也为$1/\sqrt{2}$。在认定“观察”被进行之前，必须认为两部分（在正常的向前时间描述中）“共存”。在进行观察的那一瞬间，这些共存的选择将自己分解成实际的选择——一种或另一种选择各自具有由这些幅度平方得到的概率，即$(1/\sqrt{2})^2=1/2$。当进行观察时，光子被反射或透射的概率的确都是一半。

让我们看看如何把这些应用于我们实际的实验中去。假定光子发射

时L都记录下来。光子波函数在镜子处分解，它到达P的幅度为1/$\sqrt{2}$ ，这样光电管记录到或没有记录到的概率各为一半。光子波函数的另一半到达实验室墙壁的A点（图8.3），其幅度又是1/$\sqrt{2}$ 。如果P没有记录到，那么必须认为光子打到墙上的A点去。假定我们放一个光电管在A点上，只要P没有记录到，它就记录到——假定L的确记录了光子的发射——只要P记录到，它就没有记录到。在这种意义上讲，没有必要在A处放置光电管。我们可以推断，只要看L和P就可以知道A处的光电管要做什么。

必须清楚如何进行量子力学的计算。我们提如下的问题：

“假定L记录到，P记录到的概率为多少？”

其答案是，我们注意到光子通过LMP路径的幅度为1/$\sqrt{2}$ ，通过LMA路径的幅度为1/$\sqrt{2}$ 。在取平方后我们求得它到达P和A的概率各为1/2。所以，我们这问题的量子力学答案为“一半。”这的确是我们实验中得到的答案。

我们可以同样地利用怪异的“时间反演”步骤来得到同一答案。假设我们注意到P有了记录。我们考虑光子的一个时间向后的波函数，在这里假定光子最终到达P。我们沿时间相反方向追踪光子，则光子退回去直到它到达镜子M。波函数在这一点分叉，它的1/$\sqrt{2}$ 幅度到达电灯泡L，而它的1/$\sqrt{2}$ 幅度受到M的反射达到实验室墙的另一点上，亦即图8.3上的B点。我们在取平方后又得到两种可能性各一半的概率。但是我们必须仔细地留心，这些概率所回答的是什么问题。它们是这两个问题，“假定L记录到，P记录到的概率是多少？”和前面一样，这更怪异的问题是：“假定光子从墙壁的B点射出，P点记录到的概率是多少？”

在某一种意义上，我们可以认为两种答案在实验上都是“正确的”，虽然第二个（从墙上发射）是一种推断，而不是一系列实际的实验结果！然而，这些问题中没有任何一种是我们原先问过时间反演问题。那就是：

“假定P记录到，则L记录到的概率为多少”。

我们注意到，对这个问题正确的实验答案根本不是“一半”，而是“一。”如果光电管的确记录到，则光子肯定是从灯泡而不是实验室墙

壁出来！在我们时间反演问题的例子下，量子力学计算给了我们完全错误的答案！

这一事实的含义在于量子力学R部分的规则不能适用于这种时间反演的问题中，在一个已知将来态的基础上，如果我们希望计算过去态的概率，并试图采用简单地取量子力学幅度平方模的标准的R步骤，则会得到完全错误的答案。这个步骤只有在过去的态的基础上来计算未来态的概率时才可行——它在这里极其有效！基于这些，我认为它很清楚地表明了步骤R不能是时间对称的（并且，顺便提及，所以它也不能从时间对称的步骤U中推导出来）。

许多人也许会认为这种与时间对称的矛盾是由于热力学第二定律暗中隐藏在论证之中，引入了由幅度求平方步骤所未描述的附加的时间非对称性。的确，任何能够实行R步骤的物理测量仪器必须牵涉到“热力学的不可逆性”——这样，只要进行测量熵就增加。我认为第二定律很可能以一种非常根本的方式牵扯到测量过程中。而且，使诸如上述（理想化的）量子力学实验，包括全部有关的测量记录整个操作进行时间反演，似乎没有多少物理意义。我不关心一个实验的实际时间反演时人们能进展多远，我只关心这个由取幅度平方模得到正确概率的了不起的量子力学步骤之适用性。这种简单的步骤不需要任何其他关于系统的知识就能应用在未来方向上，这真是令人惊叹。这的确是理论的一部分，人们不能影响这些概率，量子理论概率是完全随机的！然而，如果人们试图把这些步骤在过去的方向应用（亦即进行回溯而不是预见），则就完全错了。不管用多少借口来解释为何幅度平方步骤不能正确地应用于过去方向，但事实总是事实，它不适用。在未来的方向上根本不需要这些借口！正如在实际应用中那样，步骤R就不是时间对称的！

霍金盒子：和外尔曲率假设的一个关联

读者不会怀疑我们所说到的这些，但是它们和WCH或CQG有什么相干呢？是的，第二定律，正如现在有效的那样，很可能是步骤R的一个部分。但是，空间——时间奇点或量子引力对这些连续地“时时刻刻”发生的态矢量减缩能有任何觉察得到的作用吗？为了表述这一问题，我想描述一个奇异的“理想实验”，这原先是史蒂芬·霍金提出的，虽然他原先的意图并不包括我在这里的目的。

想象一个极其巨大的密封盒子。其墙壁是完全反射的，并且把一切影响都阻挡住。没有物质，包括任何电磁信号、中微子或其他任何东西能穿过它。任何从外面或里面撞到上面的东西都被反射回去。甚至引力效应也被禁止通过。不存在任何可用于建造这种墙的物质。没人能在实际上进行我就要描述的“实验”。（正如我们将要看到的，也没有人愿意去实现！）但这不是关键。在一个理想实验中人们努力从虚拟的实验中纯粹用头脑进行考虑以揭示一般的原理。技术困难只要对所考虑的一般原则没有影响，则可不予理会。（回忆一下第6章中关于薛定谔猫的讨论。）在我们的情形下，为此实验目的建造墙的困难被认为纯粹是“技术性的”，可以不予理睬。

盒子内装有大量的某种物质。何种物质并非关键。我们只关心它的总质量M，它应是非常大的，以及容纳它的盒子之巨大体积V。我们利用这个造价非常昂贵的盒子以及其无趣的内容做什么呢？这实验是可以想象到的最枯燥的实验。我们将永远不去碰它！

我们所关心的问题是该盒子内容的最终命运。根据热力学第二定律，它的熵要增加到最大值，这时物质达到了“热平衡”。如果此后相对简短地偏离热平衡的“起伏”暂时不出现的话，则不会有什么太多的事要发生。我们假定在这种情形下，M足够大并且V具有相当的值（非常大，但不是过大），使得达到“热平衡”时，部分物质坍缩成一个黑洞，只有一点物质和辐射在环绕着它——构成了一个（非常冷的）所谓的“热库”，黑洞就浸在这一个热库中。为了确定起见，我们可以选取M为太阳系的质量，V为银河系的尺度！则“热库”的温度仅比绝对零度大约高10-7度。

为了更清楚地理解这种平衡和这些起伏的性质，我们回忆一下在第5章和第7章提到的相空间的概念。它和熵的定义关系紧密。图8.4简单地画出了霍金盒子内容的整个相空间P。我们记得，相空间是一个巨大维数的空间，它上面的每一点代表我们考虑的系统全部可能的态——这里系统是盒子的内容。这样，P的每一点记录了盒子中所有粒子的位置和动量以及有关盒子的空间——时间几何所有必需的信息。图8.4中右边的（P的）子区域B代表全部所有盒子里有一黑洞的态（包括所有多于一个黑洞的情形），而左边的区域A代表所有没有黑洞的态。我们设想子区域A和B应按照熵的精确定义所要求的那种“粗粒化”被进一步分割成更小的区域（参阅图7.3，395页），但是这里我们不关心其细节。

在这一阶段我们所需要注意的是，这些区域中最大的一个代表了和一个黑洞共存的热平衡，这是B的主要部分，而A的主要部分（小一些）是代表没有黑洞的，呈现热平衡的区域。

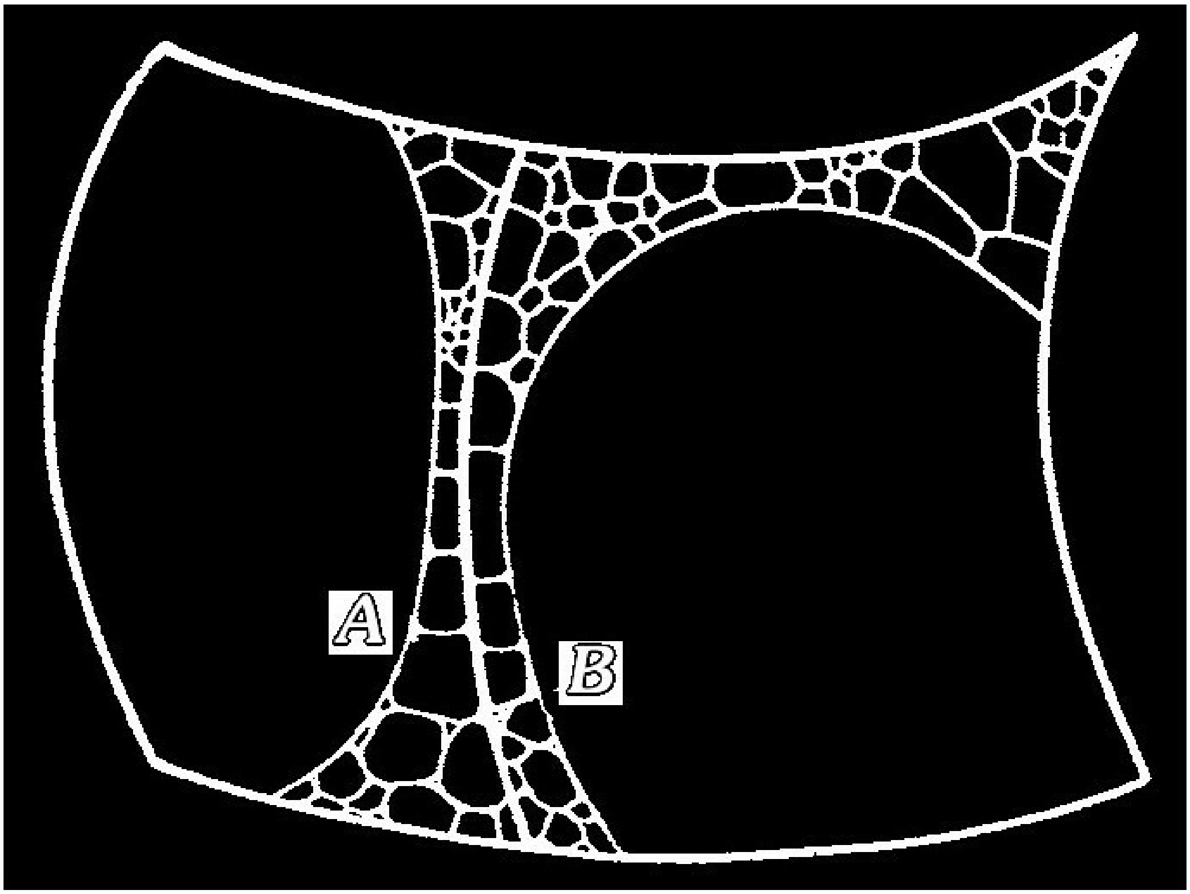


图8.4　霍金盒子的相空间P。区域A相应于盒子中没有黑洞的情形，而区域B相应于盒子中有一个（或多个）黑洞的情形

我们记得，在任何相空间中存在一个箭头（矢量）的场，它代表着物理系统的时间演化（参阅第5章229页以及图5.11）。这样，为了看下一时刻会发生什么，我就简单地跟随着P中的箭头（图8.5）。有些箭头会从区域A穿到区域B去。这种情形发生于物质因引力坍缩而形成黑洞之时。这些箭头是否会从区域B穿回到区域A去呢？是的，这是可能的，正如早先提到过的（参阅第435页，448页），只有当我们考虑了霍金蒸发的现象后才可能。根据严格的广义相对论的经典理论，黑洞只能吞没东西。霍金（1975）在考虑了量子力学的效应后，能够在量子水平上向我们展示出，无论如何黑洞必须依照霍金辐射过程发射出东西来。（这是由于“虚粒子产生”的量子过程而引起的。粒子和反粒子从真空中不断连续地产生出来，通常在其产生后立即相互湮没，不留下任何痕

迹。但存在一个黑洞时，在还没来得及湮没时，它就“吞没”了这对粒子中的一个，而它的配偶就会从黑洞逃走。此逃走的粒子构成了霍金辐射。）通常情形下，霍金辐射的确是非常小的。但在热平衡状态时，黑洞在霍金辐射中丧失的能量大小刚好和吞下其他碰巧在黑洞所在处“热库”附近徘徊的“热粒子”所获得的能量相平衡。由于“起伏”黑洞或许会非常偶然地辐射得太多或吞下得太少而失去能量。在损失能量时，它损失质量（由于爱因斯坦公式$E=mc^2$），并根据制约霍金辐射的规则，它变得更热一些。当起伏足够大时，非常非常偶然地，黑洞甚至可能进入剧烈变动的状态，它变得越来越热，失去越来越多的能量，变得越来越小，直到最终在一个（假定的）激烈的爆炸中完全消失！这情形的发生（假定在盒子中没有其他的黑洞），就对应于在我们的相空间P中从区域B过渡到A，所以确实存在从B到A的箭头！

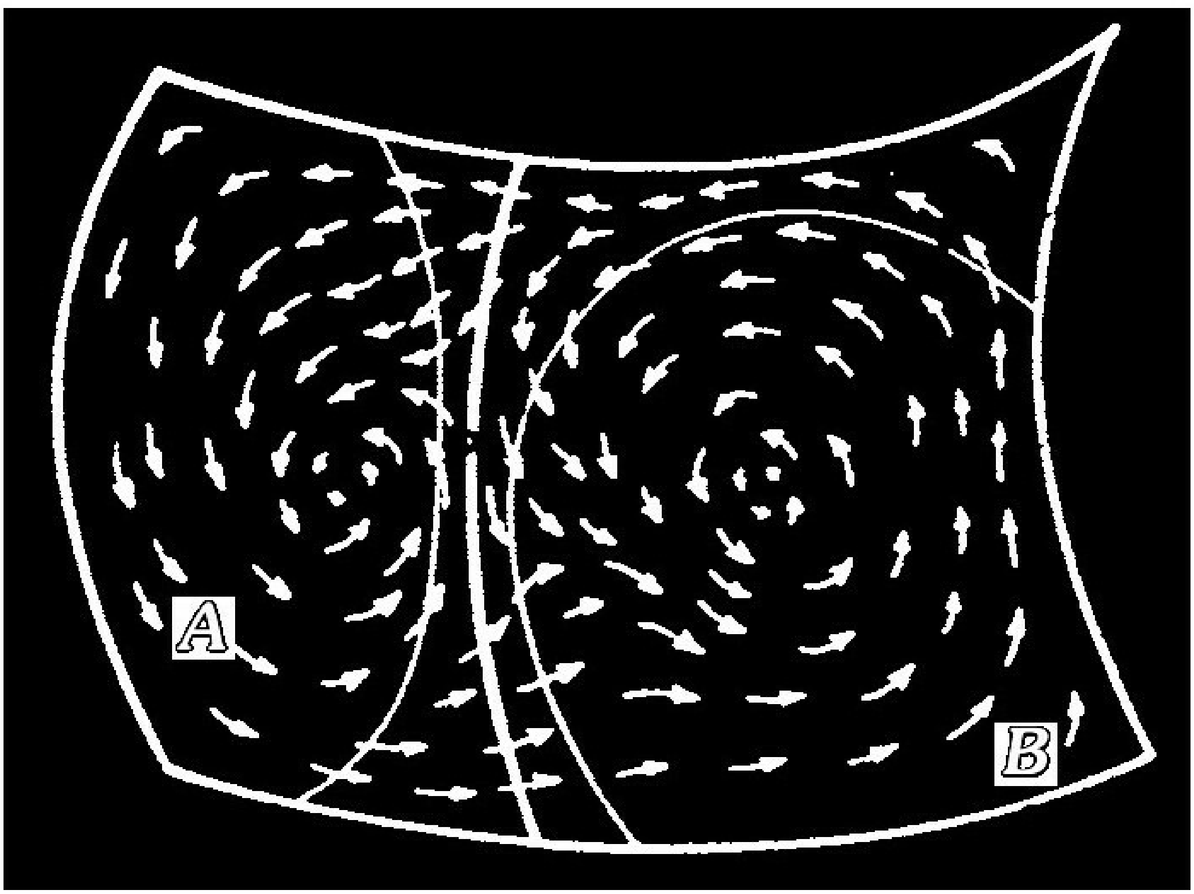


图8.5　霍金盒子内容“哈密顿流”（与图5.11相比较）。从A向B穿过的流线代表黑洞的坍缩；而从B到A的流线表明黑洞因霍金蒸发而消失

在这里我应当评论一下“起伏”是什么含义。回顾一下我们在上一章

考虑过的粗粒化的区域。属于一个区域的相空间的点被（客观上）认为相互之间是“不可区分的”。因为随着时间的推进，我们随着箭头进入越来越大的区域，所以熵就增加。最终，相空间的点停留在最大的区域中，也即相应于热平衡（最大熵）。然而，这只到某种程度为止是对的。如果人们等待足够长的时间，相空间的点会最终地跑到一个更小的区域里，而熵会相应地减少。在通常情况下它不会长久（相对而言）待在那种状态，而熵又会很快地上升，它在相空间内又进入更大的区域中。这就是伴随着熵暂时降低的起伏。熵通常不会下落太多。但是一个大的起伏会非常非常偶然地发生，而熵会降低得很多——也许会在某一较长的时间间隔中保持低值。

为了经由霍金辐射过程从区域B到区域A，这种东西是我们所需要的。因为箭头于B和A之间要穿越很小的区域，所以需要非常大的起伏。类似地，当相空间点在A的主要区域时（代表没有黑洞的热平衡），要花很长的时间才能产生引力坍缩，从而该点运动到B去。这里大的起伏又是需要的。（热辐射不容易遭受到引力坍缩！）

究竟从A到B的箭头和从B到A的箭头哪种更多，或者是一样多呢？这对我们来说是个重要的问题。换种方式来提问，在自然中由热粒子的引力坍缩形成黑洞，和由霍金辐射来排除黑洞，哪种过程更“容易”些？或者是同等“困难”？严格地讲，我们并非关心箭头的“多寡”，而是相空间体积的流率。把相空间想象成充满了某种（高维的！）不可压缩的流体。箭头代表流体的流动。回忆在第5章描述过的刘维尔定理。刘维尔定理断言，相空间体积被流线维持着，也就是说，相空间流体的确是不可压缩的！刘维尔定理似乎告诉我们，从A到B和从B到A的流量必须相等。因为相空间“流体”是不可压缩的，不能在任何一边累积起来。这样看来，从热辐射产生黑洞正如消灭它一样地“困难”！

这的确是霍金自己的结论，虽然他是基于某种不同的考虑而得到这个观点。霍金的主要论点是，所有牵涉到此问题中的基本物理都是时间对称的（广义相对论、热力学、量子力学的标准的幺正过程），所以如果我们把钟往后倒转，我们就应得到和向前走一样的答案。这归结于很简单地把在P中的所有箭头方向反转。从这个论证的确得出，从A到B和从B到A应有同样多的箭头，只要区域B的时间反演仍为区域B（而且同样，A的时间反演还是A）。这条件归结为霍金一个鲜明的设想，黑洞与其时间反演，即白洞在物理学上其实是一模一样的！他的推论是，应

用时间对称物理，热平衡态必须也是时间对称的。我不想在这里对这种奇异可能性详细讨论。霍金的思想是，在一定程度上量子力学的霍金辐射可被看作经典的物质被黑洞“吞没”的时间反演。虽然他的建议极为天才，但遇到了严重的理论困难，我不相信这能行得通。

这一建议和我这里提出的观念无论如何不能相协调，我论证过，由于外尔曲率假设，黑洞必须存在而白洞是被禁止的！WCH把时间不对称引进讨论，而霍金没有考虑到这一点。必须指出，由于黑洞及其空间——时间奇性的确是关于霍金盒子中发生事件非常重要的一部分，这里一定需要牵涉到制约这些奇点行为的未知物理。霍金认为未知物理必须是时间对称的量子引力理论，而我断言它必须是时间不对称的CQG！我声称，CQG的主要含义之一应是WCH（由此导出热力学第二定律的众所周知形式），所以我们要弄清WCH对我们这个问题的含义。

让我们看看纳入WCH会如何影响讨论P中“不可压缩流体”的流动。黑洞奇点在空间——时间中的效应是吸收并消灭所有撞到上面的物质。对我们现在的目的而言更重要的是，它消灭信息！这一效应在中是某些流线合并到一起（图8.6）。两种原先不同的态，只要把将它们区别开来的信息消灭后就会变成同一个态。但流线在中合并到一起，我们就实质上违反了刘维尔定理。我们的“流体”不再是不可压缩的，而是在区域B内被连续地湮没！

我们现在似乎陷入了麻烦。如果“流体”在区域中连续被消灭，那么从A到B的流线就会比从B到A的更多——这样产生黑洞比消灭黑洞更为“容易”！现在若不是“流体”从区域A流出比流入的更多，则这的确是有意义的。区域A没有黑洞，白洞可能性已被WCH排除掉——所以刘维尔定理在区域A应该能完美成立！然而，现在我们似乎需要某种在区域A“产生流体”的手段以补充在区域B的损失。哪些机制可以增加流线的数量呢？我们所需要的是同一个态有时能多于一个结果（亦即流线的分叉）。在将来物理系统的演化上，这类不确定性具有量子理论的“味道”——R部分。R在某一意义上能否是WCH的“硬币的另一面”呢？WCH引起了流线在B内的合并，量子力学步骤R使流线分叉。我要宣称，正是量子力学客观的态减缩（R）引起流线分叉，并由此准确地补偿了因WCH引起的流线合并（图8.6）！

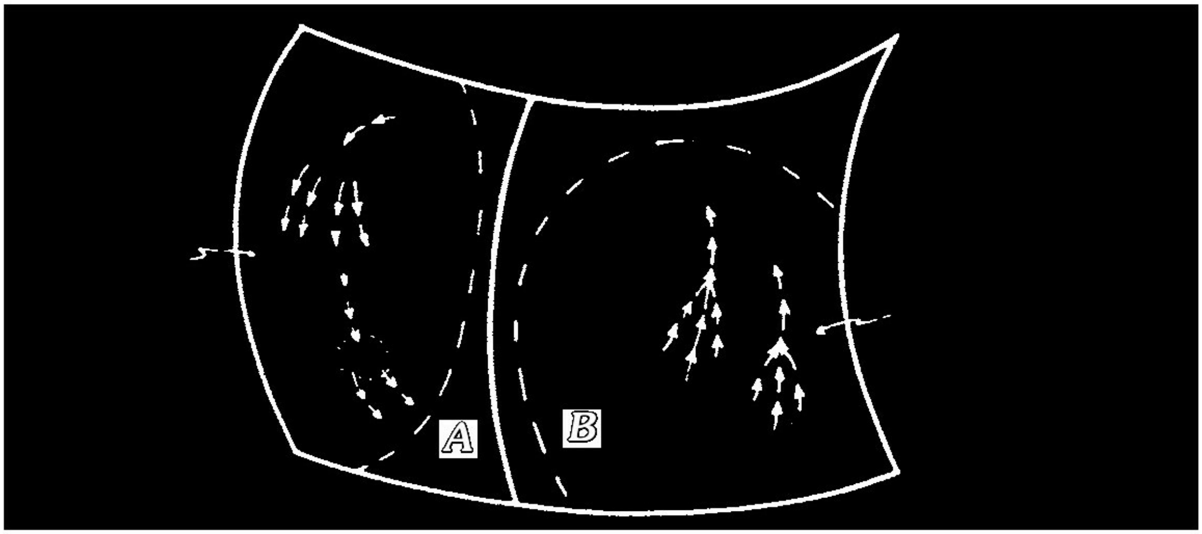


图8.6　在区域B中，由于黑洞奇点处的信息丧失，流线应该合并到一起。这是否被量子过程R（尤其是区域A中）流线的产生所平衡呢？

为了使这样的分叉发生，我们必须让R时间不对称，正如我们已经在上述的灯泡、光电管和半镀银镜子实验中看到。在灯泡发射出一个光子后，最终有两个（等概率的）选择：或是该光子打到光电管上并被它记录，或打到墙上的A处而光电管没有记录到。该实验的相空间中，我们有一根代表光子发射的流线，它分叉成两条：一条描述光电管被点燃的情形，而另一条是没有点燃的情形。这是真正的分叉，因为只允许一个输入，而却有两个可能的输出。人们也许必须考虑的另一输入是光子从墙上的B处发射出来，这时就有了两个输入和两个输出。但是，这另一选择由于它和热力学第二定律，也就是在向过去方向演化追溯时被最后表达成WCH的观点，不相协调而被排除掉。

我必须反复说明，我所表达的观点的确不是“传统的”——尽管我一点也不清楚，一位“传统的”物理学家为解决此问题有何高见。（我怀疑他们之中很少人认真地考虑过这些问题！）我当然听到过许多不同的观点。例如，时时总有一些物理学家提议，霍金辐射永远不会使一个黑洞完全消失，而某一很小的“金块”将永存下来。（所以，按照这种观点，从B到A没有流线！）这对我的论证影响很小（而实际上还会加强它）。人们还可假设相空间体积实质为无限大来逃避我的结论，但是这和有关黑洞熵某些基本思想相左，也和一个封闭（量子）系统之中相空间的性质相左。而我听到的其他在技术上逃避我结论的方法就更不能令人满意了。有一个反对观点显得稍微认真些，即在实际建造霍金盒子时需要太大的理想条件，在假定它可被造出时违背了某些原则。我本人对

此并不肯定，但倾向于相信，所需的理想条件的确是可以容忍的！

最后，我承认我掩饰了一个要点。在开始讨论时，假定我们有一经典的相空间——而刘维尔定理适合于经典物理。但是霍金辐射的量子现象必须予以考虑。（量子理论对于P的有限维数以及有限体积是必需的）。正如我们在第6章看到，相空间的量子版本为希尔伯特空间，所以在整个讨论中我们应当使用希尔伯特空间，而不是相空间。在希尔伯特空间中也存在类似的刘维尔定理。这是由时间演化U的“幺正”性质引起的。我的整个论证也许能按照希尔伯特空间，而不是经典相空间来表述，但是很难了解，如何用这种方法来讨论牵涉到黑洞时空间——时间几何的经典现象。我自己的观点是，既非希尔伯特空间也非经典相空间适用于正确的理论。人们必须利用某种迄今尚未发现的处于两者之间的数学空间。根据此观点，我的论证只能认为是处于启发性的水平上，它仅仅是建议性的，而非结论性的。尽管这样，它为WCH和R根本上相互连接，并因此为R必须是量子引力效应的想法提供有力的实例。

重述我的结论：我提出量子力学的态矢量减缩的确是WCH的另一面。根据这一观点，我们所寻求的“正确量子引力理论”（CQG）两个重要含义为WCH和R。WCH的效应为相空间中流线的合并，而R的效应刚好是补偿流线的散开。两个过程都和热力学定律紧密相关。

注意，流线的合并完全发生在区域B中，而流线的散开可在A或者B中发生。我们记得A代表黑洞的不存在，所以态矢量减缩的确在黑洞不存在时可以发生。很清楚，为了R起作用（正如在我们刚才考虑的光子实验中），不必要在实验室中有一个黑洞。我们在这里只关心在可能发生的事情中一般整体的平衡。按照我所表达的观点，只不过是说，在某一阶段形成黑洞（并因此消灭信息）的可能性必须被量子理论中不决定性所平衡！

态矢量何时减缩

假设在前面论证的基础上，接受态矢量的减缩也许最终为引力现象。R和引力的关系能解释得更显明吗？在这观点的基础上，一个态矢量的坍缩实际上应在何时发生呢？

我应首先指出，甚至在量子引力理论的“更传统的”方法中，在合并

广义相对论原理和量子理论规则时，存在某种严重的技术困难。这些规则（首先在薛定谔方程的表达式中，动量被重新解释为对位置取微分的方法步骤，参阅369页）根本不能顺应于弯曲空间——时间几何的观点。我本人的观点是只要引进“相当”大的空间——时间曲率，则量子线性叠加的规则就失效。在此处不同态的可能选择的复幅度叠加，正是被实际，也就是实在发生的，可能选择的概率权重所取代。

所谓“相当大的”曲率是何含义呢？我是指引入的曲率测度达到水平大约为一个引力子 [6] 或更大的尺度。（回忆一下，根据量子理论的规则，电磁场被量子化成单独的称为“光子”的单位。当场被分解成为它单独的频率，频率v的部分只能以整数个光子出现，每一光子具有hv的能量。类似的规则应可适用于引力场。）根据量子理论，一个引力子应是被允许的最小曲率的单位。其想法是，只要到达这个水平，依据U过程的线性叠加的通常规则在应用到引力子时就被修正，而某种时间不对称的“非线性不稳定性”就出现。在这一阶段，其中一种选择就脱颖而出，该系统就“跌跌撞撞”地落到这种选择之上，而不再以复数线性叠加不同选择的形式永远存在下去。也许选择的结果是由机遇造成，也许在这后面还有更深刻的东西。但是现在，现实已成为这种或那种选择。R步骤就这么得以完成。

请注意，根据这个思想，步骤R以一种完全客观的方式自动发生，和任何人为的干涉无关。其想法是，“单引力子”水平必须安宁地处于原子、分子等通常量子理论的线性规则U成立的“量子水平”以及我们日常经验的“经典水平”之间。单引力子水平的“尺度”多大呢？应强调的是，这实在不应当是物理上的大小的问题；它更应是质量和能量分布的问题。我们看到只要不牵涉到太多的能量，量子干涉的效应可在大距离上发生。（回忆在322页描述的光子自干涉以及367页克劳塞和阿斯匹克特的EPR实验。）质量的量子引力的特征尺度为所谓的普朗克质量（大约估计）：

$m_p = 10^{-5}$ 克。

这似乎比人们希望的大很多，由于质量比这小很多的物体，诸如灰尘，以经典方式行为就能直接感受到。（质量m_p 比虱子的质量小些。）然而，我认为单引力子的标准不可以就这么生硬地使用。我试图弄得更显明一些，但就在我写作的现在，关于如何准确使用单引力子的

标准，还有许多模糊之处。

首先，让我们考虑一个观察粒子非常直接的方式，也就是利用威尔逊云室。此处有一个小室充满了刚好处于就要凝聚成液滴的蒸汽。当一个快速运动的带电粒子，譬如刚从位于小室外的放射性原子衰变而产生的，进入这个小室时，在它通过蒸汽的路途中，会使近处的某些原子电离（也就是，由于失去电子而带电）。这种离化了的原子成为蒸汽凝聚成小液滴的中心。我们以这种方法得到实验者可直接观察的小液滴的轨迹（图8.7）。

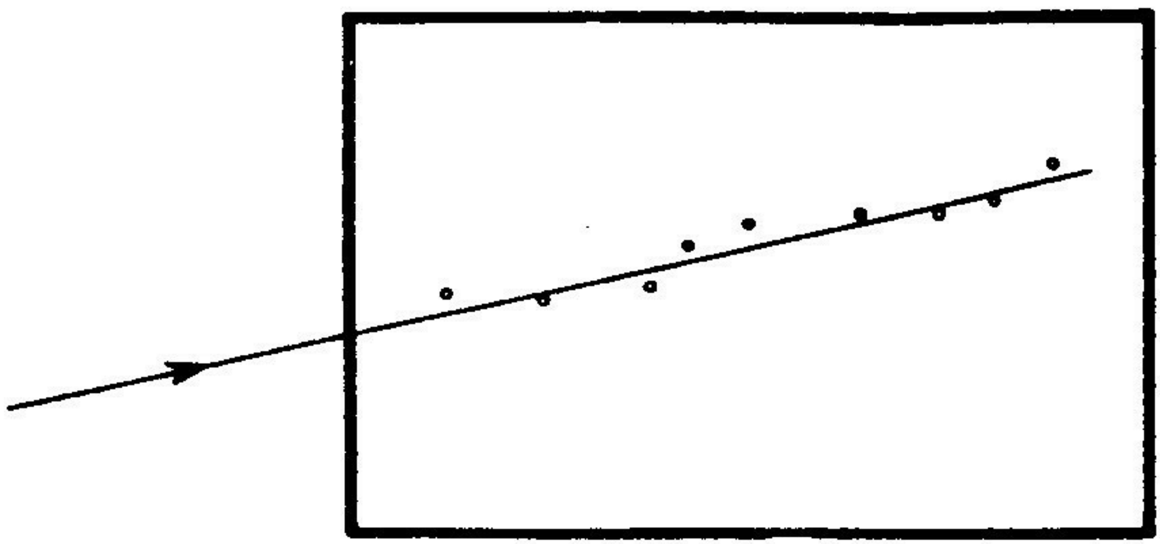


图8.7　一个带电粒子进入威尔逊云室并引起一串液滴的凝聚

现在，如何利用量子力学对此作描述呢？在我们放射性原子衰变的时刻，它发射出一个粒子。但是，该粒子可往许多不同的方向飞离。在这一方向有一幅度，在那一方向又有一幅度，在其他每一方向都有一幅度，所有这些都在量子线性叠加上同时发生。这些叠加的不同选择的总体组成了从衰变原子出发的球面波：被发射出的粒子的波函数。当每一可能的粒子轨道进入云室，它就和一串电离的原子相关联，每一个原子成为蒸汽凝聚的中心。所有这些不同可能的离化原子串也须在量子线性叠加中共存，所以我们现在有大量不同的凝聚水滴串的线性叠加。在某一个阶段，当按照步骤R取复幅度加权的平方模后，这个复数的量子线性叠加变成了实在的不同选择的实概率加权集合。这些选择只有一个在经验的物理世界中实现，而这一个特殊选择即被实验者所观测到。我于是根据这一种观点提议，只要不同选择的引力场差别达到一个引力子水

平，这一个阶段就发生。

这在什么时候发生呢？根据非常粗略的计算 [7] ，如果只有一个完全均匀的球滴，则当水滴长大到大约1/100m p也就是1克的1/10^7 时，即达到了单引力子的阶段。在此计算中存有许多不确定性（包括某些原则上的困难），而且为保险起见，值取得稍大一些，但其结果并非完全不合理。人们期望以后将会得到更精密的结果，并能处理整串液滴而不仅仅是一粒液滴。当人们考虑水滴是由大量的小原子组成而非整体均匀的，这事实也许会导致某些重大的差别。另一方面，“单引力子”标准本身在数学上必须变得更加精密。

我已在上面的情况考虑了在一个量子过程（放射性原子衰变）中的实际观察。量子效应被放大到这种程度，此时不同的量子选择产生不同的、直接可观察的不同选择。我的看法是，即使当这种显明的放大不存在时，R还可客观地发生。假定一个粒子不进入云室，而是直接进入到一个装满气体（或流体）的大盒子，气体（或流体）密度使得它肯定和粒子碰撞，或者该粒子扰乱大量的气体原子。让我们仅仅考虑粒子的两种不同选择，将其当成原先复数线性叠加的部分；或者它根本不进入该盒子，或者它会沿着特定的路径进入，并且掠飞过一些气体原子。在第二种情形下，气体原子会以巨大的速度跑开，如果该粒子没有进入，气体原子就不会这样地行为。它会继续碰撞并掠飞过其余的原子。这两个原子中的每一个都以一种前所未有的方式飞走，并很快地引起气体原子的连锁运动。如果原先粒子没有进入盒子，这一切都不会发生（图8.8）。在第二种情形下，不用很长时间，气体中的每一个原子实际上都被这个运动所扰动。

现在让我们想想应该如何用量子力学来描述。最初，在复线性叠加中只需要考虑原先粒子的不同位置——将其当作粒子的波函数部分。但是在很短的时间后，所有气体原子都被涉及，考虑该粒子可能采取的两个途径的复线性叠加，一个途径进入盒子，而另一个途径不进入盒子。标准的量子力学坚持，我们要把这种叠加推广到气体中的所有原子：我们必须把两种状态叠加，一种状态下的所有原子都从另一种状态下的位置移动开。现在考虑所有原子总体的引力场的差别。尽管气体的整体分布在叠加的两种状态下实际上是一样的（而且整体引力场也如此），如果我们把一个场减去另一个场，便得到一个（高度振荡的）差场。在现在我所关心的意义上，该差别也许是“重要的”——也就是说这个差场很

容易超过单引力子水平。只要达到这水平，态矢量减缩就会发生：在系统的实际状态中，或者是粒子进入，或者没有进入盒子。复线性叠加被归结为统计的加权的不同选择，其中只有一种真正发生。

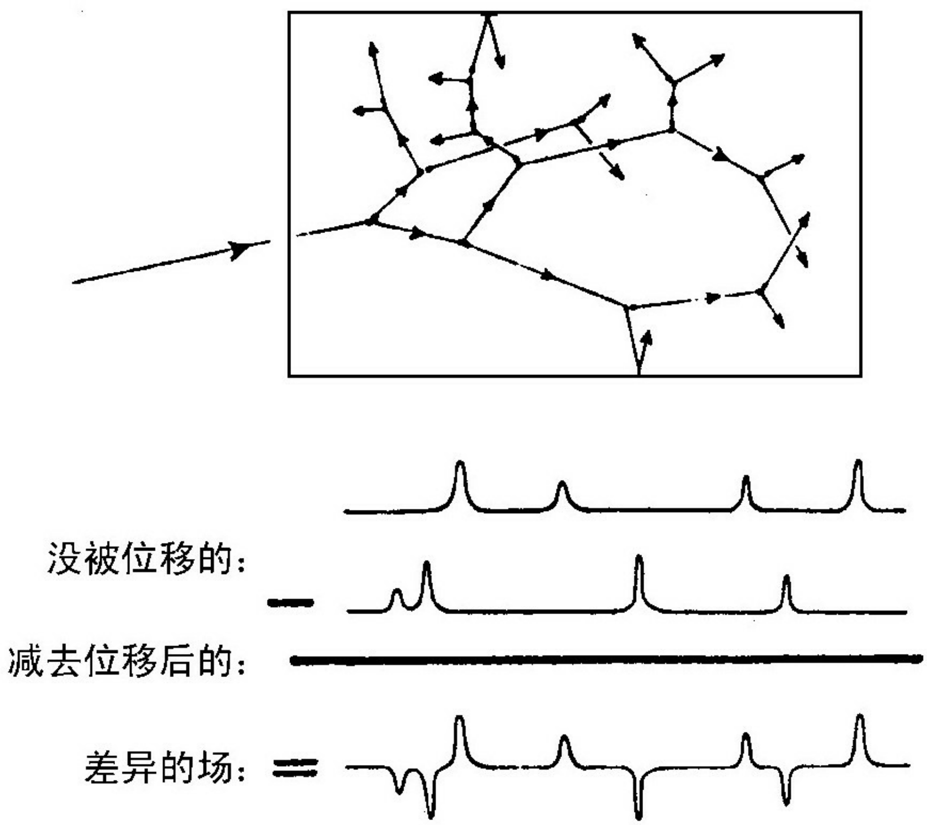


图8.8　如果一个粒子进入到某种气体的大盒子中，则实际上气体的每一个原子很快就都会受到扰动。粒子进入和粒子没进入的量子线性叠加会牵涉到描写气体粒子的这两种形态引力场的不同空间——时间几何线性叠加。这两种几何的差别什么时候达到单引力子的水平呢？

在前例中，我把云室当作提供量子力学观察的一种方法。依我看来，其他类型的观察（照相底片、火花室等）似乎也可利用“单引力子判据”来处理，正如我在上述的气体盒子情况所用的方法。为了解这一步骤的细节还有许多事要做。

到此为止，这只是一个观念种子，我相信它是非常需要的新理论[8]。我坚信，任何完全满意的方案必须牵涉到空间——时间的某种激进的新观念，也许是某种本质上非定域的描述[9]。这个信念最吸引人的理由来自于EPR类型的实验（参阅358页，367页）。这种实验中，在屋子

一个角落的观察（这里指光电管记录）会在另一角落引起态矢量的瞬息减缩。建立一个和相对论精神一致又完全客观的态矢量减缩理论是一个深远的挑战。因为在相对论中，“同时性”不是一个恰当的概念，它依赖于某些观察者的运动。我的意见是，眼下的物理实在的图像，尤其在和时间性质的关系中，将要受到巨大的，也许甚至迄今为止比相对论和量子力学所引起的都要大的冲击。

我们应回到原先的问题上来。所有这一切如何与制约我们大脑行为的物理相关联呢？它和我们的思维以及情感有何关系呢？为了回答这类问题，就必须首先考察我们大脑的实际构造。我在后面将要回到我认为是基本问题上来。当我们有意识地思维或感觉时，会牵涉到何种新的物理行为？

第9章　真实头脑和模型头脑

头脑实际上是什么样子的

在我们的脑袋中有一个控制我们动作并使我们了解周围世界的非常了不起的结构。正如阿伦·图灵有一次说过[1]，它和一碗凉粥再相像不过了！非常难以想象，具有如此平淡无奇外观的东西怎么能创造这么多的奇迹。然而，更周密的考察揭示，头脑的结构极其错综复杂（图9.1）。盘旋在顶上的（最像粥样的）巨大部分称为大脑。它很清楚地从中间分成左边和右边两个大脑半球。它的前面和后面不那么清楚地分成了额叶和三片其他的叶：顶叶、颞叶和视叶。再下去，头脑后面小很多而有点像球形——或像两团羊毛球——的部分是小脑。里面深处藏在大脑下面有些奇怪而显得很复杂的不周结构：脑桥和髓质（包括我们以后要关心的网状结构），它们由脑干、丘脑、下视丘、海马、胼胝体和其他许多命名奇怪的古怪构造组成。

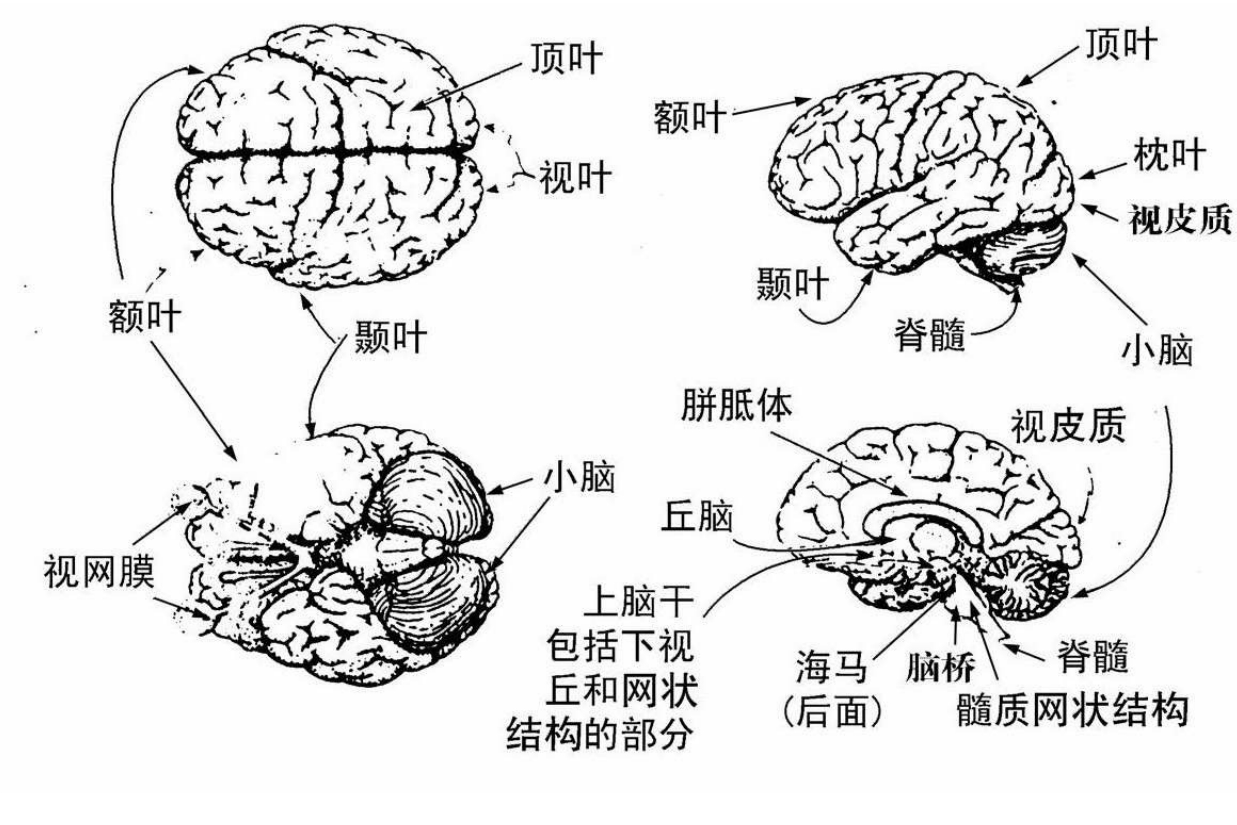


图9.1　人脑：上、边、下、剖视图

人类感到最为骄傲的即是大脑，不仅因为它是人脑中最大的部分，而且作为整体而言，人的大脑在比例上比其他任何动物的都大。（人的小脑也比大多数其他动物的大）。大脑和小脑具有比较薄的灰色物质外表面，以及具有白色物质的更大的内部区域。人们把这些灰色物质的区域分别称为大脑皮质和小脑皮质。灰色物质正是实行不同种类计算任务的地方，而白色物质是由很长的神经纤维所组成，负责从头脑中一个部分传递信号到另一部分。

大脑皮质不同的部位和非常特别的功能相关联。视皮质处于视叶内，在大脑的正后方，它与图像的接收和解释有关。至少对人类而言，大自然会选择这个区域来解释从头部正前方的眼睛传来的信号是很奇怪的！但是，大自然还有比这更古怪的行为。正如大脑的右半球几乎完全关联身体左边，而大脑的左半球关联身体的右半部分。这样，实质上所有的神经在进入或离开大脑时都要从一边穿到另一边。对于视皮质的情形，右边并不和左眼相关，而是和两只眼睛的左边视野相关。类似地，左边的视皮质和两只眼睛的右边视野相关。这表明，从每只眼睛视网膜的右边来的信号必须进入右边的视皮质（记住视网膜上的影像是上下颠倒的），而从每只眼睛的视网膜左边来的信号必须进入左边的视皮质（图9.2）。这样在右边的视皮质中形成了一个左边视野的一个界限清楚的映射，而在左边视皮质中形成了右边视野的另一个映射。

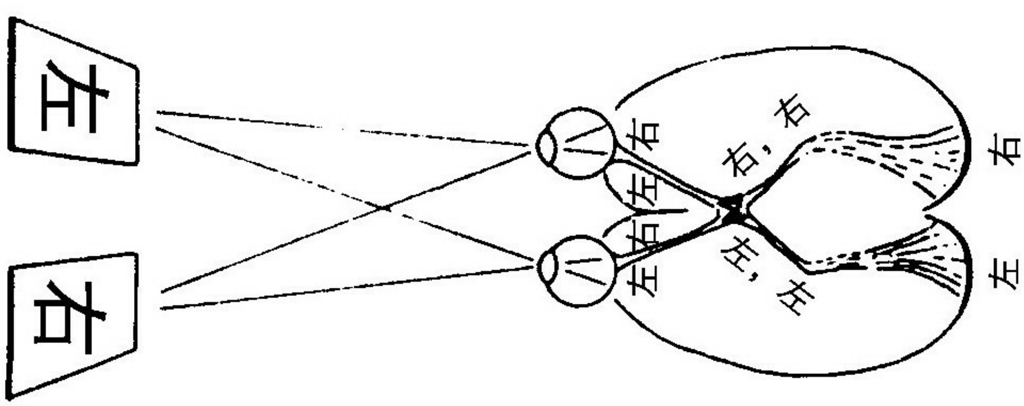


图9.2　两只眼睛的左边视野被映射到右边的视皮质上去，而右边视野被映射到左边的视皮质上去。（下视图；注意视网膜中的像是颠倒的。）

从耳朵来的信号又是以这种古怪方式穿越到大脑的相反一边。右边

的听觉皮质（为右边颞颥叶的一部分）主要是处理在左边接收到的声音，而左边的听觉皮质处理从右边来的声音。嗅觉似乎是这个一般规则的例外情形。位于大脑前面的右边嗅觉皮质（位于额叶内——对感觉区域而言，这很特殊）主要是应付来自右鼻孔的气味，而左边的是处理左鼻孔的气味。

触觉和称之为触觉皮质的顶叶区域相关。这个区域刚好发生于额叶和顶叶分开的地方。在身体表面和触觉皮质的区域之间有一种非常特别的对应。有时按照所谓的"触觉侏儒"的图示法来表示这一种对应，也就是图9.3中沿着触觉皮质躺着的变形人体。右边触觉皮质处理左边的身体，而左边的处理右边的身体。在额叶和顶叶的交界处正前方，有一个称为运动皮质的额叶对应区域。这和激发身体不同部位的动作有关，在身体不同的肌肉和运动皮质的不同区域之间又有一个非常特别的对应。我们用一个"运动侏儒"来描绘这一种对应，正如图9.4所表示的那样。右边的运动皮质控制身体的左边，而左边的运动皮质控制右边。

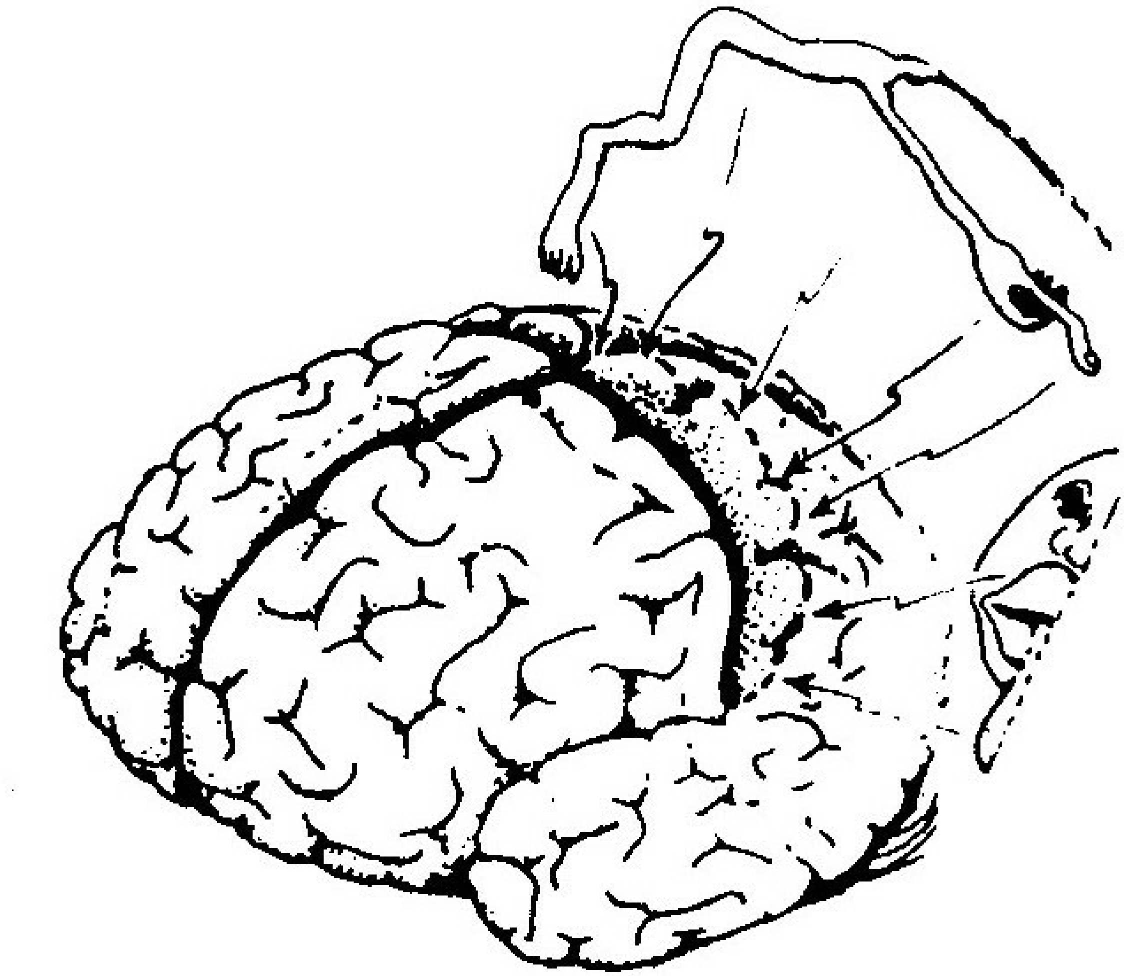


图9.3“触觉侏儒”图解额叶和顶叶分界后面的大脑部分，它和身体不同部位的触觉有最直接的相关

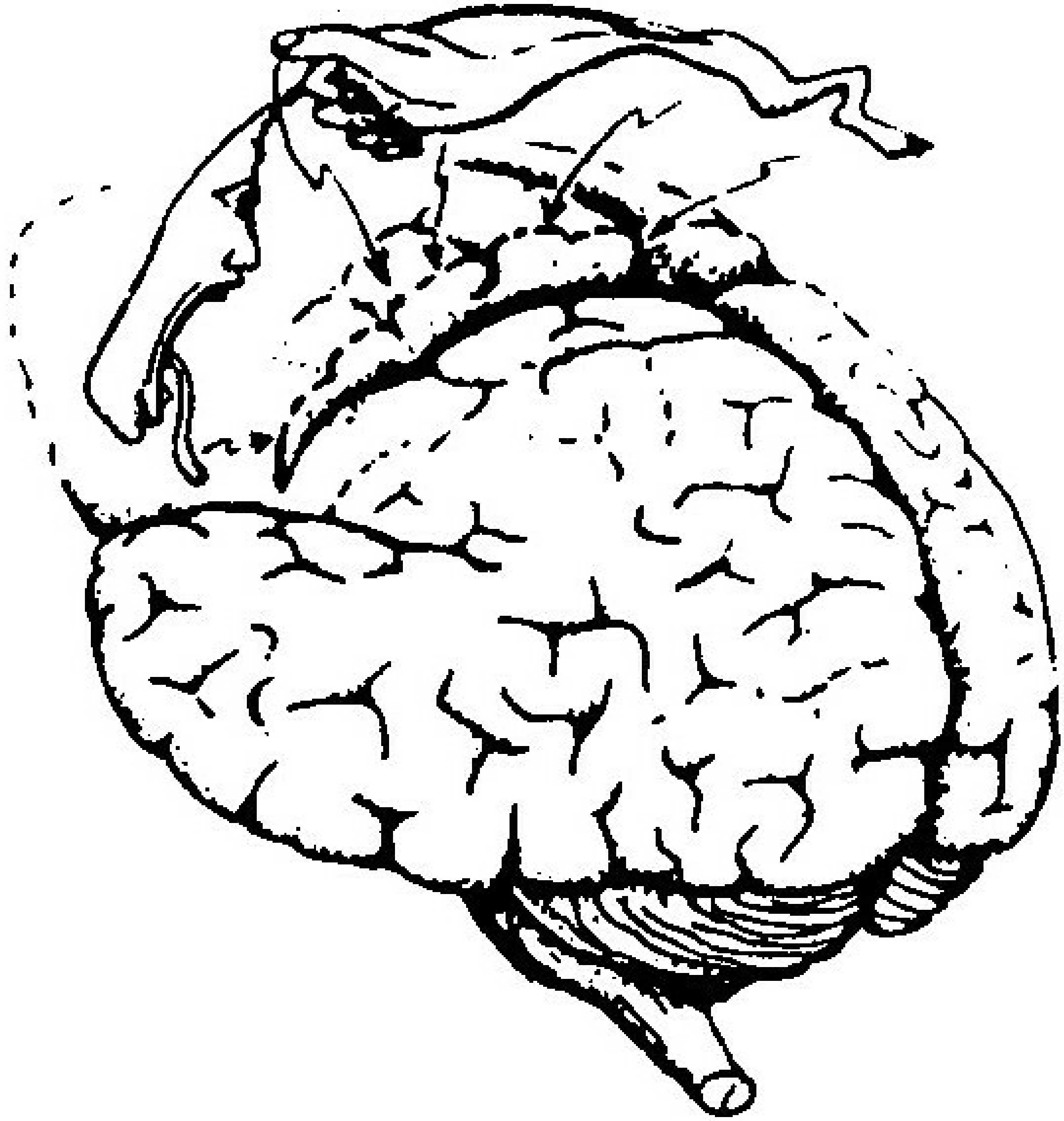


图9.4“运动侏儒”图解额叶和顶叶分界前面的大脑部分，它最直接激发身体不同部位的动作

刚才提到的大脑皮质的区域（视觉的、听觉的、嗅觉的、触觉的和运动的）被称作首位的，这是由于它们和大脑的输入和输出有最直接的关联。和这些首位区域邻近的是大脑皮质第二位区域，它处理更微妙更复杂的抽象感觉（图9.5）。从视觉、听觉和触觉皮质接收到的感觉信息在相关的第二位区域加工，而第二位运动区域和预想的动作相关，这些动作由首位运动皮质翻译成实际肌肉运动更特定的指令。（由于嗅觉皮质的行为不同，我们又对其所知甚少，所以先不予考虑。）大脑皮质的其余区域称作第三位区域（或缔合皮质）。头脑最抽象最复杂的活动

大多在这些第三位区域进行。在一定程度上，头脑正是在这些区域，连同它们的周围，使不同感觉区域来的信息以非常复杂的方式相互交流并接受分析，留下记忆，建立外界的图像，构想并衡量一般计划以及理解和表达语言。

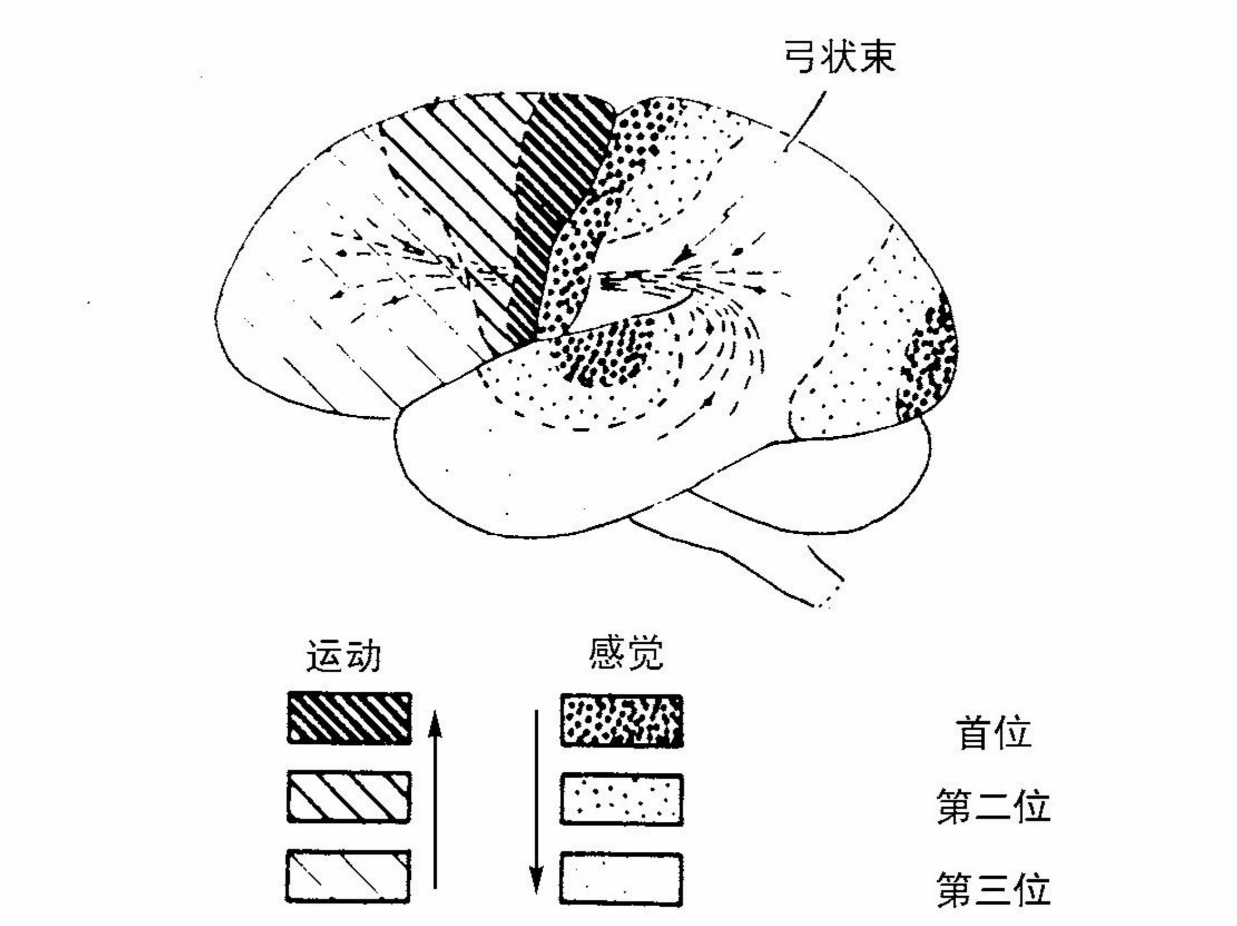


图9.5　大脑作用的大致划分。外部的感觉数据进入首位的感觉区域，在第二位和第三位的感觉区域逐步加工到越来越复杂的程度，转移到第三位的运动区域，最后改善成为在首位运动区域的特殊的动作指令

语言是特别有趣的，这是由于它通常被认为是人类智慧所特有的能力。很古怪的是（至少在绝大多数用右手的人和大部分用左手的人），语言中心主要是在头脑的左边。主要的区域是伯洛卡区域，这是处于额叶的后下部，还有另一区称作温尼克区域，它处于颞颥叶的后上部（图9.6）。伯洛卡区域是用来形成句子，而温尼克区域是理解语言；损害伯洛卡区域会减少讲话能力，但不影响理解；而损害温尼克区域，则讲话仍然流利而没有什么内容。称作弓状纤维束的神经束把这两个区域连接起来。当它受到损坏时，理解力没有受到伤害而且言语仍然流利，但

是不能把所理解到的讲出来。

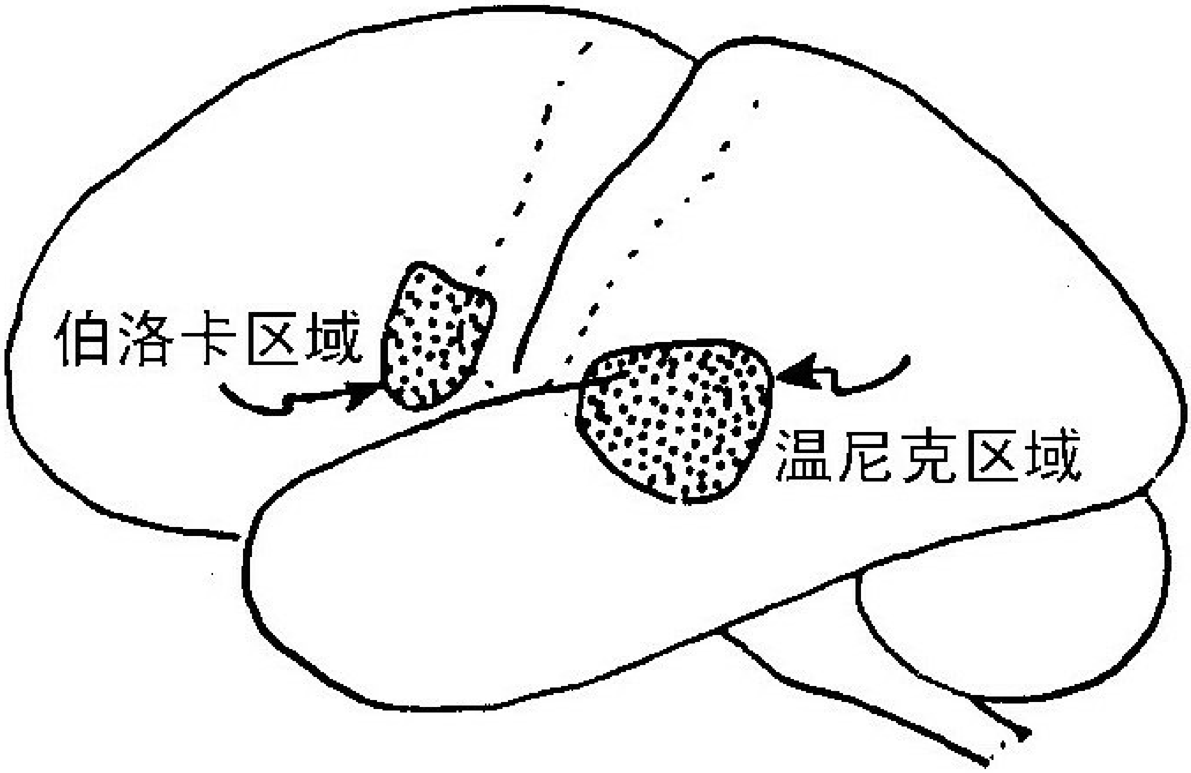


图9.6　通常只在左边：温尼克区域和理解语言有关，而伯洛卡区域和表达语言有关

我们现在对大脑的作用有了一个非常粗略的图像。头脑的输入来自于视觉、听觉、味觉和其他的信号，这些信号首先要在大脑的首位部分，主要是后叶片（顶叶、颞叶和枕叶）记录下来。头脑的输出表现于激发身体的动作，主要由大脑额叶的主要部分完成。在它们之间进行了某些加工过程。一般来讲，大脑活动是从后叶的首位部分开始，当分析输入数据时，就进入到第二位部分；然后当这些数据被完全理解时，就进入了后叶的第三位部分（也就是语言在温尼克区域被理解）。弓状纤维束，也就是上面提及的神经纤维束，但现在是在头脑的两边，把这处理过的信息带到额叶，在额叶的第三位区域形成一般行动的计划（例如，在伯洛卡区域形成语言）。这些一般的行动计划在第二位运动区域被翻译成关于身体运作更明确的概念，并且头脑的活动最后转到首位运动皮质，在那里信号最终被发送到身体不同组的肌肉去（而且经常一下子送到好几组肌肉去）。

呈现在我们面前的似乎是一台超等计算仪器的图像。强人工智能支

持者（参阅第1章等）会坚持道，我们在这里有了一台算式电脑的极致范例，实质上就等于是图灵机。该电脑有输入（如同图灵机左边的输入磁带）和输出（如同机器右边的输出磁带），以及在中间实行各种复杂的计算。当然，头脑还能独立进行特殊感觉输入以外的活动。这在人们思维、计算或对过去的回忆沉思时发生。对于强人工智能的支持者而言，这些头脑的活动只不过是进一步的算式活动。只要这种内部活动达到足够复杂的程度，则他们就可以设想引起了“知觉”的现象。

然而，我们不应该太快接受现成的解释。上面所表述大脑活动的一般图像只不过是非常粗糙的图像。首先，甚至连视觉接受也不像我在上面表述得那么一清二楚。人们发现，皮质中有几个不同的（虽然是更小的）区域进行视野的映射，这显然是带有其他不同的目的。（我们视觉的知觉似乎和它们不同。）似乎还有其他辅助的感觉和运动区域分散在大脑皮质（例如，后叶中不同的点可引起眼睛的运动）。

我在上面的描述中，甚至没有提到头脑中大脑以外其他部分的作用。例如，小脑的作用是什么？显然它是负责身体的准确定位和控制，身体动作的时机、平衡和精巧。想象舞蹈家熟练的艺术，职业网球运动员轻松的准确性，赛车手闪电般的控制以及音乐家或画家手的自如动作；再想象羚羊优美的跳跃和猫的躲藏。没有小脑就不可能有这样的准确性，所有动作都会变得笨拙。事情似乎是这样的，当一个人在学习新技巧（例如学走路或学开车）时，刚开始对于每一个动作都要仔细想好，大脑起着控制作用；但是一旦掌握了技巧而成为“第二天性”，小脑就取代了它。而且，人们对这种经验都很熟悉，如果一个人在想已经掌握的技巧中的一个动作，则他的容易控制会暂时失去。在想动作时似乎涉及重新引进大脑的控制，虽然在此以后获得动作的灵活性，但现时熟练和精确的小脑作用却损失了。这种描述无疑过于简单，不过使我们对小脑的作用有些大概的了解[64]。

在我早先描述大脑作用时，没有提到头脑的其他部分是误导的。例如，海马在记录长期（永久）记忆时有着不可或缺的作用。实际的记忆被储存在大脑皮质的某处（也许同时在许多处）。头脑可以用其他方式保持短期的印象；它可保留印象若干分钟甚至几个小时（也许把它们“记在心里”）。但是为了使这种印象在不再受注意时还能被回忆起，就必须以永恒的方式存在那儿，正是这个原因海马非常重要。（损害海马会导致可怕的后果，一旦新的记忆不再引起病人的注意，即不再保

留。）胼胝体是左右两个大脑半球相互通信的地方。（我们在后面将会看到把胼胝体切除后的严重后果。）下视丘是快乐、愤怒、害怕、沮丧和渴望的情感所在处，而且它传递情感在精神和身体方面的发泄。在下视丘和大脑不同部分之间存在连续流通的信号。丘脑的作用是重要的加工中心和转换站，它把从外面世界来的许多神经输入传到大脑皮质。网状结构负责头脑整体或不同部位一般状态的警戒和知觉。还有许多神经线路连接这些以及许多其他极其重要的区域。

上面的叙述只提供了头脑一些重要部分的样品。在结束这一节之前，我应该对于头脑的整体组织再讲一些，它的不同部分被分类成三个区域，从脊柱开始往上，按顺序是后脑（或菱脑）、中脑和前脑。人们可以在早期发育的胚胎中找到这三个区域，在脊柱的上端按照这个顺序成为三个肿胀出现。最顶点的那个即是正在发育的前脑，发芽成两个球状的肿胀，一边一个，后来它们变成大脑的两半球。完全发育好的前脑包括头脑许多重要部分，不仅是大脑，还有胼胝体、丘脑、上视丘、海马以及许多其他部分。小脑是后脑的一部分。网状结构在中脑和后脑各有一部分。在进化发展的意义上，前脑是“最新的”，而后脑是“最古老的”。

我这个简要的速写虽然在多方面是不足够的，希望仍能给读者一点人脑像什么样子和一般情形下它做什么的印象。迄今，我几乎还未触及意识的中心问题。让我们在下面讨论这个问题。

意识栖息在何处

有关头脑的状态和意识现象的关系，人们表达了许多不同的观点。对于具有这么明显重要性的现象只有极少的共识。然而，下面这一点是很清楚的，即头脑的所有部分不是同等地牵涉到意识的呈现。例如上面所暗示的，小脑似乎比大脑更可被视为一台“自动机”。仿佛小脑控制的动作几乎是不必思考就自动进行的。人们能够有意识地决定从一处走到另一处，但他不会时时想到为控制运动所必须详细计划的肌肉动作。对无意识的反射行为亦是如此，譬如把自己的手从热火炉上移开的动作，可能根本不是由头脑而是由脊柱上部分传递的。人们至少从这些很容易推理出，意识现象和大脑的作用比和小脑或脊髓的作用更有关系。

另一方面，十分不清楚的是，大脑是否总能发觉自己的活动呢？例如我在上面所描述的，人们在正常行走时，并不意识到自己肌肉和四肢的细节活动。这种活动的控制主要来自于小脑（头脑的其他部分和脊髓予以帮助），大脑的首位运动区域似乎也参与控制。此外，首位感觉区域也是一样：例如，人们可能不知道走路的时候在脚底的压力变化，但是触皮质的相应区域仍然继续受到刺激。

杰出的美国——加拿大神经外科医生怀尔德·彭费尔德（大部分人脑运动和感觉区域内的精细映射是他在20世纪40—50年代做出的）论断说，一个人的知觉不只和大脑活动相关联。基于对许多有意识的病人进行脑手术的经验，他提出，主要由丘脑和中脑组成被他称为上脑干的区域（参阅*Penfeld and Jasper 1947*），在相当意义上应被视为是“意识所在处”，尽管他在心里曾经基本认同网状结构作为“意识所在处”。彭费尔德断言，上脑干和大脑处于联络状态，只要脑干的这个区域和大脑皮质的适当区域（也就是和任何具体感觉、思维、记忆相关的特别区域）处于直接联络的状态，或者那时候动作被有意识地发觉或唤起，则“意识知觉”或“有意识的意志行为”就会产生。他指出，例如，当他刺激病人引起右臂动作的运动皮质的区域（而且右臂的确会动作），这不会使病人想要运动其右臂。（的确，病人甚至可能伸出其左臂去阻止右臂的动作，正如在彼得·塞勒著名的斯特任奇拉夫博士的电影描写中那样！）彭费尔德指出，动作的欲望也许和丘脑的关系比和大脑皮质的关系更大。他的观点认为意识是上脑干活动的呈现，但是除此以外必须还要有东西被意识。这样，不仅是脑干，而且此时正和上脑干联络的大脑皮质的某些部位都被涉及，它们的活动代表了意识的主体（感觉印象或记忆）或客体（意志行为）。

其他神经生理学家还指出，特别是网状结构可能被认为是意识的“所在处”，如果这样的处所的确存在的话。网状结构毕竟是负责头脑一般的警觉状态（*Moruzzi and Magoun 1949*）。如果它受到损害，其后果就是变成无意识。只要头脑处于清醒的意识状态，那网状结构便是活跃的；如果头脑不处于这种状态，网状结构就不活跃。在网状结构的活动和人们通常认为“意识的”状态之间的确有种清楚的关联。然而，情况由于以下事实变得复杂起来。在梦境里，在当时人们确实晓得自己在做梦的这个意义上讲是“知觉的”。然而，这时在正常情况下网状结构的活跃部分仿佛不处于活跃状态。要把意识这样的荣耀地位归于网状结构的人还要忧虑一件事，用进化的术语讲，它是头脑中非常古老的部分。如

果成为有意识的全部所需只是一个活跃的网状结构，那么青蛙、蜥蜴甚至鳕鱼都是有意识的！

我个人认为上面这论断不是很有说服力。我们有何证据说蜥蜴和鳕鱼不具备某种低程度的意识呢？我们有什么权利像一些人那样宣称，人类是我们行星上仅有的被赐予“知觉”的实际能力的居民呢？在地球生物中，难道我们是唯一可能“有意识”的吗？我表示怀疑。虽然青蛙和蜥蜴，特别是鳕鱼，在我凝视着它们时，它们并没使我十分确信“那里有某一个生灵”也在看着我。但是当我看到狗或猫，尤其是动物园的猩猩和猴子在看着我时，我有很深的“意识存在”的印象。我不要求它们像我这样感觉，甚至也不要求它们感觉有多少深度。我不坚持它们在任何强烈意义上是“自我知觉的”（尽管我猜想自我知觉的某一因素能够存在[65]）。我所坚持的是，它们有时候至少有感觉！至于做梦的状态，我自己愿意接受的是，存在某种形式的知觉，但是它被认定为低水平的。如果部分网状结构以某种方式单独负责知觉，那么它们必须是活跃的，即使在做梦状态时其活跃的程度很低。

另一种观点（*O’Keefe 1985*）仿佛认为，海马的行为和意识状态的关系更大。正如我早先评论的，海马对于长久记忆的记录十分重要。有理由可以认为，永久记忆记录和意识相关联。如果真如此，则海马在意识知觉的现象中的确起了主要的作用。

还有人坚持大脑皮质本身负责知觉。由于大脑是人的骄傲（虽然海豚的大脑也一样大！）而且和智力关系最密切相关的心理活动似乎是由大脑执行，那么这里肯定是人的灵魂栖息之所！这正是有些人，譬如强人工智能者观点的结论。如果“知觉”只是算法之复杂性的一个特征，或者说算式的“深度”或“微妙程度”，那么按照强人工智能的观点，由大脑皮质进行的复杂算法使该区域具有显示意识能力的呼声最高。

很多哲学家和心理学家似乎认为，人类意识是和人类语言密切相关。因而，正是多亏了我们的语言能力，人类才能得到微妙的思考能力。它正是人性的标志以及我们灵魂的表现。按照这种观点，正是语言把我们和其他动物区别开来，并提供我们剥夺它们自由以及肆意捕杀它们的借口。正是语言允许我们用哲理推究并描述我们所感觉的，这样我们可以使别人信服，我们知悉我们的客观世界和自我。正是从这一种观点出发，我们的语言被当作我们具有意识的关键因素。

现在，我们必须想起我们的语言中心（对绝大多数的人）正好在我们头脑的左边（伯洛卡和温尼克区域）。刚刚表述的观点仿佛意味着，意识只和左边而不和右边大脑皮质相关联！的确，这似乎是很多神经生理学家的看法（特别是约翰·埃克勒斯，1973年），虽然对于我这个门外汉来说，这观点确实非常古怪，我将会解释其理由。

头脑分裂实验

我将提到和这些相关的许多奇异的观察，在这些试验中病人（或动物）的胼胝体完全被割除，使得大脑皮质的两个半球不能互相联络。在人的情况下 [2] ，切除胼胝体是作为治疗手术来进行的，人们发现这是对受癫痫症之苦特别严重的病人有效的处方。罗杰·斯帕雷和他的助手在对这些病人动手术后进行了许多心理学试验。他们是这样进行的，左边和右边的视界受到完全分开的刺激，这样左半球只接受放在右边的视觉信息，而右半球只接受到左边的。如果用铅笔的画片在右边闪动，同时用杯子的画片在左边闪动，则病人会说“那是一枝铅笔”，因为只有铅笔而不是杯子被显然有语言能力的那一半头脑所感知。然而，左手会选一只盘子，而不是一张纸，去和杯子作适当的相配。左手在右半球的控制之下，右半球虽然不能讲话，却能实现某些相当复杂的人类特有的动作。的确有人建议过，认为几何思维（尤其是三维的）还有音乐通常主要是在右半球内进行的，这样就和左边的言语和分析能力相平衡。右边头脑能理解普通的名词或基本句子，并能进行非常简单的算术。

这些头脑分裂实验中最令人吃惊的是，头脑两边仿佛像两个独立的个人那样行为。每一半可分别与实验者联络。由于右半球缺乏口语的能力，它比左半球的联络更为困难，并处于更原始的程度。病人大脑的一半可和另一半以简单的方式联络，例如看着由另一半控制的手臂运动，或者听到指令的声音（像是盘子的碰撞声）。但是，即使是两边之间这么原始的联络也可由仔细控制实验室条件而消除。模糊的情绪仍可由一边传到另一边，然而这可能是由于那些没分开的部位（譬如下视丘）仍然与两边处于联络状态。

人们忍不住会问这样的问题：在我们同一身体里是否居住有两个分别意识的个体呢？这个问题曾是许多争议的主题。有人会坚持说肯定“是”的，而其他人宣称两边都不能成为单独存在的。有人会争论道，

对于两边能够有共同的情绪这个事实，可证明只含有一个单体。然而，另一种观点认为，只有左边半球代表了有意识的个体，而右边为一台自动机。认为语言是意识主要部分的人相信这一点。确实，只有左半球能令人信服地对口头问话“你有意识吗？”答复“是的！”而对右半球，正如一条狗、一只猫或一只黑猩猩，会困难得甚至不能解释组成这个问题的词，也不能够正确地口头回答。

然而，这个问题不能就这么轻易地放过。道纳尔德·威尔逊和他的合作者（*Wilson et al.1977*；*Gazzaniga, LeDoux and Wilson 1977* ）在一个更新近而引人注意的实验中考察一位代号“P.S.”的病人。在分裂手术后，只有左半球能讲话，但是两个半球都能理解语言；右半球后来也学会讲话；两个半球很明显地都是有意识的。而且，由于它们有不同的爱好和需求，所以显得有各自的意识。例如，左半球描述说它希望成为制图员，而右半球希望成为赛车手！

我本人根本不相信普遍的断言，认为平常人类语言于思维和意识是必需的。（我将在下一章指出我的一些理由。）所以，我和那些大致相信头脑分裂病人的两半各具有独立意识的人立场相同。P.S.的例子强烈地暗示，至少在这一种特殊情形下，两半的确都是有意识的。依我的意见，在这一方面P.S.和其他人的真正差别是他的右边意识能实际使其他人相信它的存在！

如果我们接受P.S.的确具有两个独立的精神，则就面临着令人惊异的情景。假定，每一分裂头脑的病人在手术之前只具有单独的意识；但在此之后就有了两个意识！这原先单独的意识被某种方式分叉了。我们还记得在第1章32页假想的旅行家，他把自己交给一台远距运送机，忽然醒过来时发现所谓“实在的”自我已到达金星。在那里，他的意识的分叉仿佛会导出佯谬。因为我们会问：“他的意识流‘实际上’是沿着什么途径？”如果你是这旅行家，哪一个被归结成“你”？远距运送机可当成科学幻想而不再予以考虑，但是在P.S.的情形下，我们的处境非常相似，而这是实在发生的！哪个P.S.的意识是手术前的P.S.？无疑许多哲学家会将此问题斥为无稽。这是由于没有决定这个争论的有效办法。每一个半球都在分享施行手术前的意识之记忆，并且无疑两者都会宣称是原来的那个人。这虽令人惊异，但本身并非佯谬。尽管如此，还有一些令人困惑的问题未被解决。

如果两个意识后来又被合并在一起，则困惑就更恶化了。现代技术要把胼胝体单独切断的神经接在一起仿佛是天方夜谭。但是人们可以摹想，在开初时改用一种比实际切除神经纤维更温和的办法。也许这些纤维可被暂时冷冻，或用某种药麻痹。我不知道这类实验是否做过，但是我认为在不太久的将来会变成可行的。可以设想，在胼胝体重新激发后，结果只会有一个意识。想象这一个意识就是你本人！对于你在过去某一时刻曾经有过两个显著不同的各自的“自我”将作何感想呢？

盲视

头脑分裂实验似乎至少指出，“意识”的“栖息处”不必是唯一的。但是还有其他实验暗示，大脑皮质的某些部分比其他部位与意识有更密切的关联。其中一种和盲视现象有关。视皮质某个区域的损害会引起相应视野的盲视。如果一个对象放在那个区域的视野，那个对象就不会被感知。相当于那个视觉区域发生了盲视，然而，一些古怪的发现指出（参阅*Weiskrantz 1987* ），事情并非这么简单。一位名叫“D.B.”的病人有些视皮质必须被除去，这使他在视野特定区域内看不见任何东西。然而，当把某物放在这个区域并要D.B.去猜是什么东西时（通常像一个十字形或圆形标志或是以某一角度倾斜的线段），他几乎可百分之一百地猜中！这个“猜测”的准确度甚至使D.B.本人都感到惊讶！他仍然坚持说，自己不能感知放在该区域的任何东西。[66]

由视网膜接受到的影像还在视皮质以外的某些头脑区域加工，一个较不清楚的区域是下部颞叶。D.B.应该是基于在下部颞叶得到的信息作“猜测”。在刺激这些区域时不能直接有意识地知觉任何东西，但是信息在哪儿，只有从D.B“猜测”的正确性中得以显现。实际上在经过一些训练后，D.B.能够得到这些区域有限量的实际知觉。

所有这些似乎表明，大脑皮质的某些地方（例如视“皮质”）比其他地方和意识知觉更有关联。但是，其他一些区域经过训练后，显然能进入直接知觉的范畴。

视皮质的信息加工

在如何处理接收到的信息方面，人们对视皮质比头脑中的任何其他部分都理解得更好；人们为了说明这个作用曾提出了不同的模型 [3] 。事实上，有些视觉信息的处理在到达视皮质之前就在视网膜本身进行。（视网膜实际上被当成头脑的一部分！）一些最先暗示视皮质如何处理信息的实验中有一个为大卫·胡帕勒和托斯腾·温塞尔赢得1981年的诺贝尔奖。他们能在实验中显示猫的视皮质某些细胞对视野中具有特定斜率的线产生反应。而其他附近的细胞对不同斜率的线产生反应。什么东西具有这个角度通常没有关系。它可以是从亮处到暗处的或从暗处到亮处的边界线，或仅仅是亮的背景中一条黑线。进行考察的特殊细胞已经把“斜率”这一特征抽象化了。而其他细胞对特定的颜色或对每只眼睛接收到的差别产生感应并由此感知所得到的深度。随着距离首位接受区域越来越远，我们发现细胞对我们所看见的东西越来越细微部分敏感。例如，当我们看图9.7时就会发觉一个完全的三角形图像；而在图上并没有画出形成三角形的线，它是由推断而来的。人们实际上发现视皮质（所谓第二位视皮质）中的细胞能记录这些推断出的直线位置！

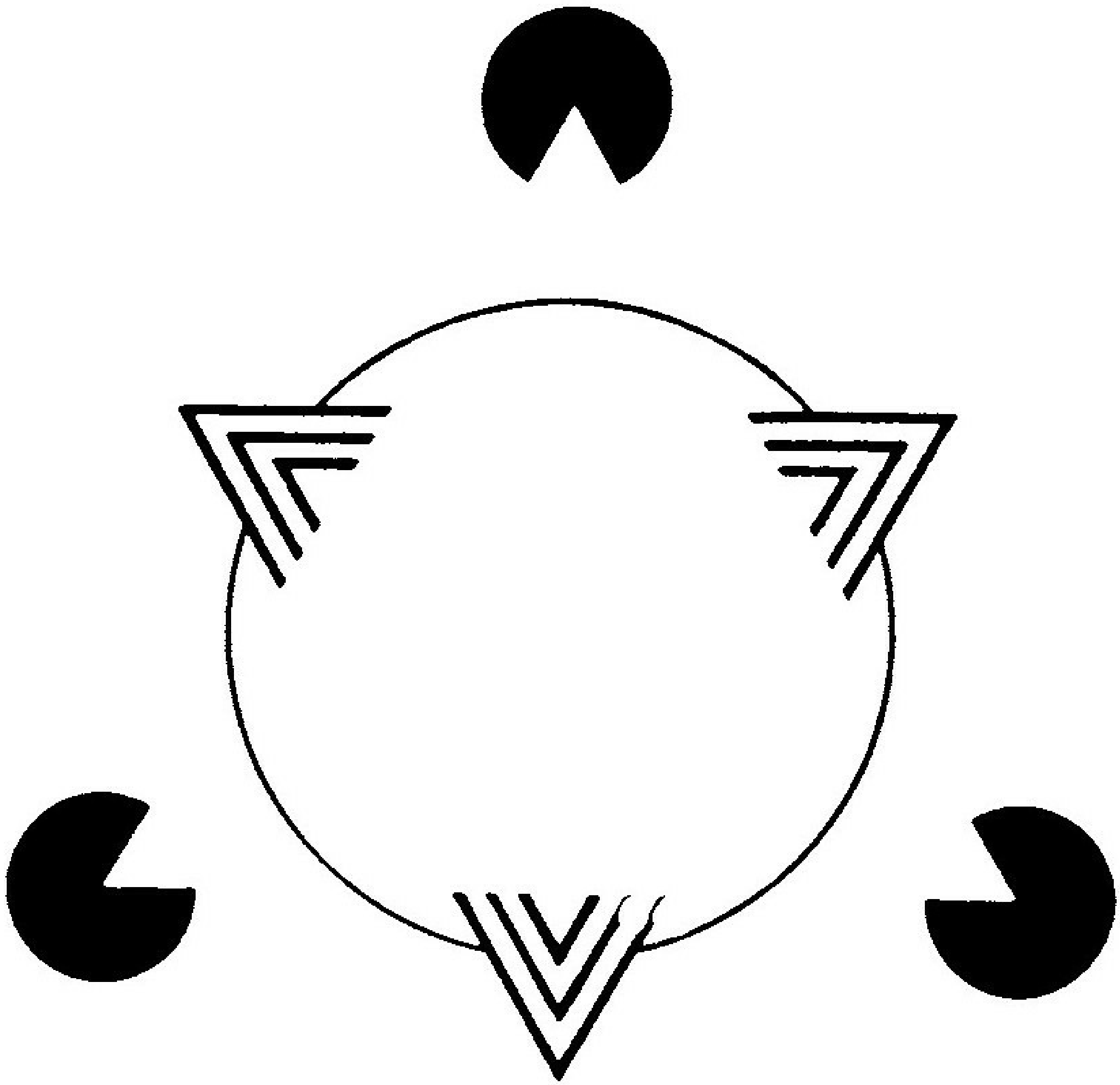


图9.7　你能看到一个放置在另一个三角形上面的，并被一个圆环嵌住的白三角形吗？虽然白三角形的边界没有完全画出来，但是头脑中有些细胞对这些看不见的，却可感知的直线有反应

在20世纪70年代初叶，文献记载宣称[4]，发现猴子视觉皮质有一个细胞，只有当一个脸的像被记录在视网膜时才反应。在这种信息的基础上，人们提出了“祖母细胞假设”。根据这种假设，头脑中存在某种细胞，它只有当头脑主人的祖母进入房间才反应！的确，最近有人发现某种细胞只对特殊的词有反应。也许这正进一步支持祖母细胞假设！

关于头脑中进行的处理过程，肯定有极多尚待研究的地方。对这些高等头脑中心如何实行其功能所知甚少。现在让我们放下这个问题，而转去注意头脑中实际产生这些奇异功能的细胞。

神经信号如何运行

头脑（以及脊柱和视网膜）所有的处理过程都是由称作神经元的多功能体细胞完成的 [5] 。让我们看看神经元是什么样子。我在图9.8中画出了一个神经元的图。它有一个苞状的主体，或有点像星形，通常比较像萝卜，这主体称为胞体，其中包含有细胞核。一根很长的神经纤维从胞体的一端延伸出来。相对于一个单独的微细胞而言，神经纤维有时的确非常长（以人类而言通常达几厘米），它称作轴突。细胞的输出信号就是沿着轴突这根“导线”传送。从轴突可发芽出许多小的分支，它要分支好几回。在这些形成的神经纤维终端可找到所谓的小突触结。胞体的另一端通常向四面八方往外分支，这就是树状的树突。输入资料正是沿着树突进入胞体。（在树突之间偶然也会出现突触结，叫作树状树突的突触。由于它们的复杂性超过其重要性，我将在讨论中省略。）

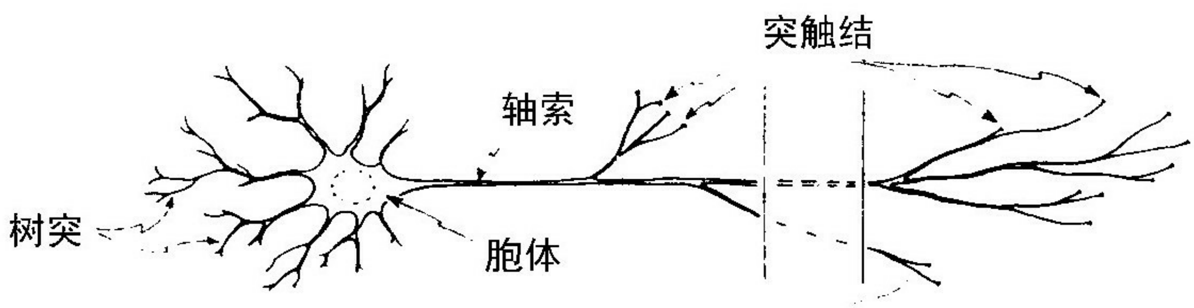


图9.8　一个神经元（通常比所标的要相对地长非常多）。不同类型的神经元在外表细节上的变化非常巨大

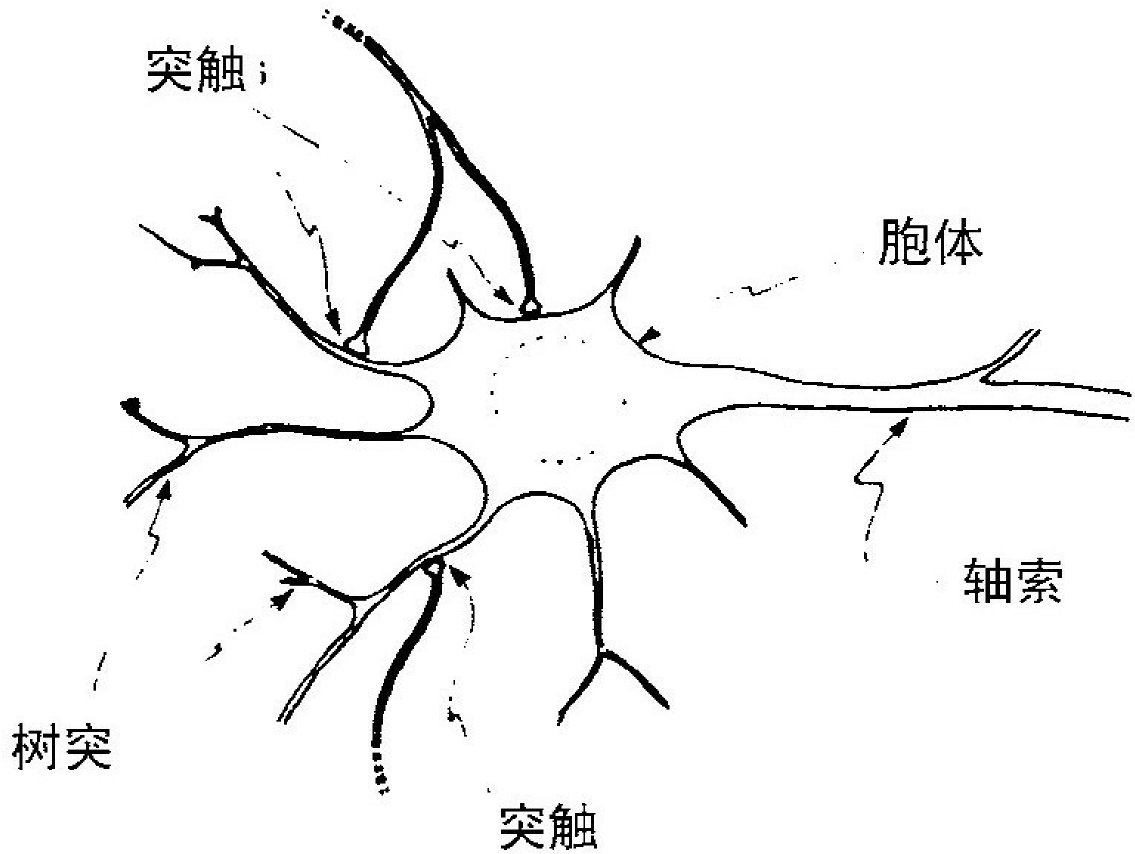


图9.9　突触：一个神经元和另一个神经元之间的结

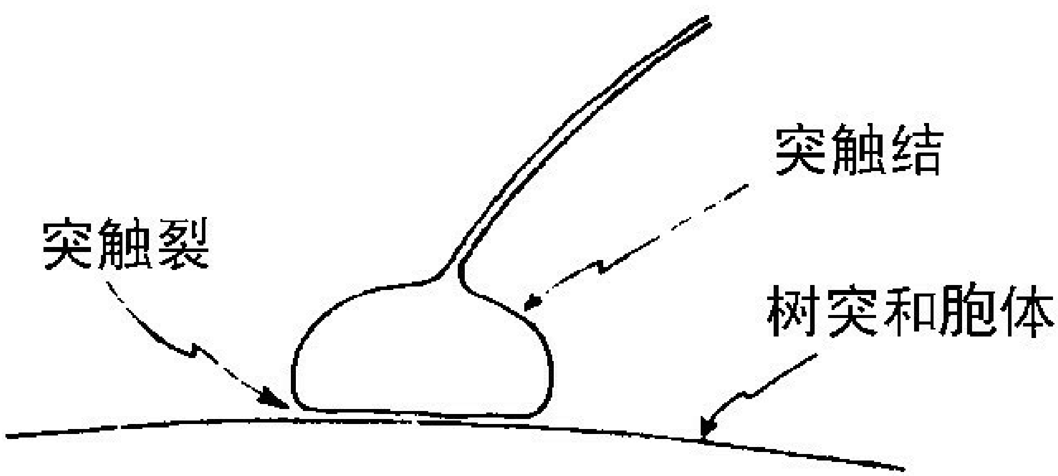


图9.10　突触的放大细节。神经传递物的化学物质流穿越过一个狭窄的缝隙

作为一个自足的单元，神经元用细胞膜把胞体、轴突、突触结、树

突和所有一切都包围起来。为了使信号从一个神经元传到另一个，它们必须设法“越过之间的障碍”。这是在叫作突触的交结处完成。一个神经元的突触结附到另一神经元的某一点上，不是在它的胞体本身就是在它的一个树突上（图9.9）。实际上，在突触结和它所附上的胞体或树突之间有一非常窄称为突触裂的缝隙（图9.10）。信号从一个神经元传达到另一个时必须穿越过它。

当信号沿着神经纤维传递并穿过突触裂时采取什么形式呢？是什么引起下一个神经元去发射信号呢？从我这个局外人看来，大自然实际采取的步骤真是非同寻常——简直使人着迷！人们也许会认为信号只不过像沿着导线的电流，但是它比这要复杂得多。

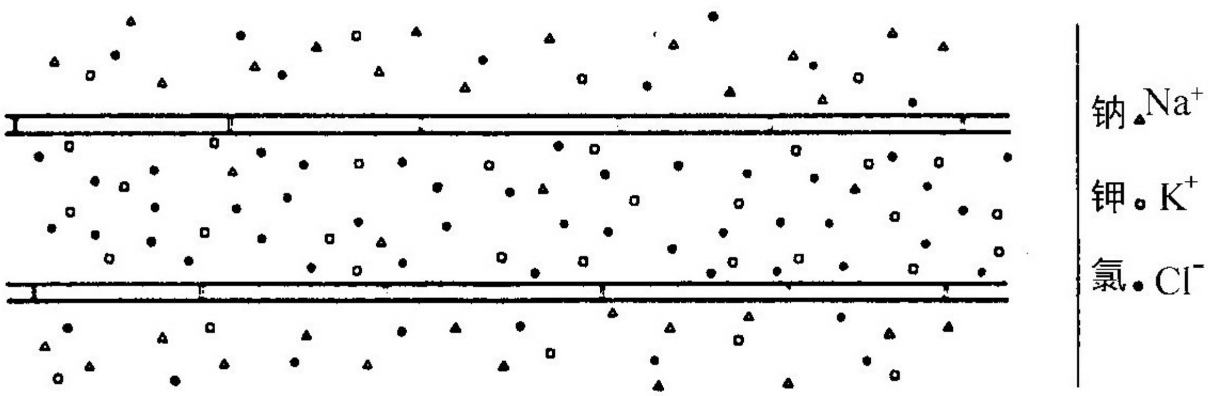


图9.11　神经纤维的图解。在静态下内部的氯离子量超过钠和钾离子的量，所以带负电；而外部刚好相反，所以带正电。内外的钠/钾平衡也不一样，在里面钾多一些，而外头的钠多一些

一根神经纤维基本上是一个圆柱管，装有普通盐（氯化钠）和氯化钾（主要是后者）的混合溶液，这样在管子里有钠、钾和氯离子（图9.11）。在外面也有离子，但是其成分比例不同，外部的钠离子比钾离子多。神经处于静态时，管子内带负电荷（也就是氯离子比钠离子加钾离子的总和还多——我们记得钠离子和钾离子是带正电的，而氯离子是带负电的），而在外部是带正电荷（也就是钠和钾比氯多）。构成圆柱表面的细胞膜有点“漏洞”，所以离子有移动穿越过它并中和电荷差的倾向。为了补充以及维持里面过量负电荷，一个“新陈代谢泵”缓慢地把钠离子通过周围的膜向外输送回去。这作用同时可维持内部钾比钠的数量更多，还有另一个新陈代谢泵（稍微小一些）把钾从外头输送到里面去，所以对内部的钾过量有贡献（虽然它的作用和保持电荷不平衡相反）。

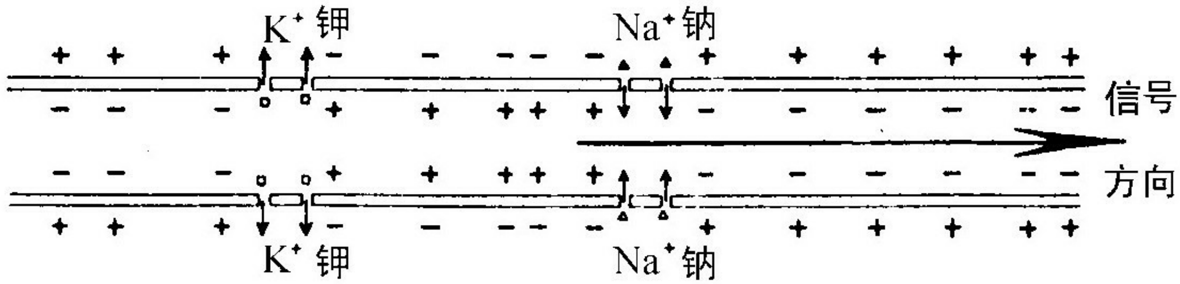


图9.12　一个神经信号是沿着纤维运动的相反带电区域。在它的前头，钠门打开让钠往里流；在它的后头，钾门打开让钾往外流。新陈代谢泵的作用是恢复原状

一个信号是沿着神经纤维移动的一个区域，其中具有相反的电荷不平衡（也就是现在内部是正的，外部是负的）（图9.12）。想象有人位于神经纤维上这种反向带电区域之前方。当这个区域靠近时，它的电场在细胞膜上打开了称为钠门的小“门”；这允许钠离子从外面往里面流回去（由电力和因浓度差引起的压力，亦即“渗透压”的结合效应）。结果使内部带正电而外部带负电。这些发生过后，构成信号的反转电荷区域即到达我们的位置。它现在促使另外一个小“门”（钾门）打开，这“门”允许钾离子从里面往外面流回去，这样开始恢复内部超量的负电荷。信号现在就通过了！最后，随着信号再次远离而去，泵缓慢而坚决的行为再次把钠离子赶到外面去，而把钾离子赶到里面去。这就恢复了神经纤维的静态，并为下一个信号做好准备。

值得注意的是，信号就是由一个沿着纤维移动的相反带电区域所构成。实际的物质（也就是离子）移动得非常小，只是在细胞膜两边穿进穿出！

这种古怪奇异的机制显然极其有效。不管是脊椎动物还是无脊椎动物都普遍利用它。但是脊椎动物有个更完善的设施，也就是神经纤维由称为髓鞘质的白色脂肪物质绝缘层所包围。（正是髓鞘质层使头脑的“白色物质”呈现其颜色。）这种绝缘使神经信号能不衰减并以高达每秒120米的相当的速度（在“转换站”之间）行进。

当一个信号到达一个突触结时，它会发射出称为神经传导物的化学物质。这种物质从它的一些树突或胞体本身的某一点通过突触裂移动到另外的神经元去。有些神经元具有一种突触结，它会发射出神经传导化学物质，趋向于促进下一个神经元躯体去“激发”，也就是开始一个沿着它的轴突传出去的新信号。这种突触称作兴奋突触。另外有一种倾向于

阻碍下一个神经元激发称作抑制突触。在任何时刻要把活跃的兴奋突触这效应全部加起来，并减去活跃的抑制突触的效应，如果净值达到某一临界值，则下一个神经元就会被激发。（兴奋突触引起下一个神经元里外之间预期的正电位差，而抑制突触引起预期的负电位差。这些电位差适当地加在一起。当附在轴突处的电位差达到临界水准时神经元就被激发，使得钾跑出来的速度不至于快到能恢复平衡。）

电脑模型

神经传导的一个重要特点是其（大部分）信号完全是“全有或全无”的现象。信号强度不变化：它要么有要么没有。这使得神经系统的行为具有类似数字电脑的特点。事实上，大量互相联结的神经元其行为与具有导线和逻辑门（下面还要讲到）的数字电脑内部作用有很多相似处。在原则上用电脑模拟特定神经元系统的行为并不困难。有一个问题就自然产生了：这是不是意味着，不管头脑的线路细节是什么样子，总是可以用电脑的功能来模拟？

为了使这个比较更清楚，我应该讲清楚究竟什么是逻辑门。在一台电脑中，我们也有所谓“全有或全无”的情形。导线不是有电流脉冲，就是没有电流脉冲。在有电流脉冲时强度总是相同的。由于每件事都被非常准确地计时，没有脉冲亦是个明确的信号，并会被电脑“注意到”。实际上，在我们使用术语“逻辑门”时，我们隐含用脉冲有和无分别来表示“真”或“伪”。这真伪与实际的真理和谬误毫无关系；只不过借以了解通常用的术语而已。让我们还把“真”（脉冲存在）写为数字“1”，“假”（没有脉冲）写为“0”，而且正如在第4章那样，可以用“&”当作“并且”（这个“陈述”表示两者都是“真”的，也就是两个论断都为1时，其答案为1），“∨”表示“或者”（这“表明”两者之一或两者都为“真”时答案为1，也就是唯有两个论断都为0时答案为0），“⇒”表示“意味着”（也就是A⇒B表示“如果A是真的，那么B是真的”，这和“或者A是假的，或者B是真的”相等），“⇔”表示“如果并且如果”（两者皆“真”或两者皆“假”），以及“～”表示“非”（如果“假”的话为“真”，如果“真”的话为“假”）。人们可把这些不同的逻辑运算列成所谓“真值表”：

$$A\&B:\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\qquad A\vee B:\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$$

$$A\Rightarrow B:\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\qquad A\Leftrightarrow B:\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

在每种情形下A标记行数（也就是A=0表示第一行，A=1表示第二行），类似地B标记列数。例如A=0，B=1表示每一个表右上方的值，而在第三表上A⇒B得到1。（作为实际逻辑的例子：断言“如果我在睡觉，那么我就快乐”肯定是可证实的；在特殊情形下，即如果我刚好醒着而且快乐时，肯定被无趣地证实。）最后，“非”逻辑门的效应简单地是：

～0=1和～1=0。

这些是逻辑门的基本类型。还有一些其他类型，但是所有那些都可由刚才提到的类型组合而成[6]。

现在，我们可否在原则上用神经元的联结建造一台电脑呢？我将指出，甚至从我们前面刚讨论过的神经元激发的非常原始的考虑而言，这的确是可能的。让我们看看如何在原则上用神经元的联结建造逻辑门。我们需要一些数字编码的新方法，因为信号不存在时不能触发任何东西。让我们（完全随意地）用双脉冲表示1（或“真”），而单脉冲表示0（或“假”），并且采取一个简单的方案，就是激发神经元的临界值总是刚好两个同时发生的兴奋脉冲。很容易建立一个“与”门（也就是“&”）。正如图9.13所示，我们可取两根输入神经纤维作为输出神经元上的仅有一对突触结的终端。（如果两者都是双脉冲，则第一次脉冲

和第二次脉冲都达到所需要的2倍脉冲的临界值，如果有一根只有一次脉冲，则两根中只有一根达到临界值。我假定脉冲被非常仔细地计时，而且是双脉冲的情形，为了确定起见，用这双脉冲中的第一次计时。）建造一个“非”门（也就是“～”）要更复杂一些，图9.14画出了一种办法。输入信号沿着一根分成两个分支的轴突进来。有一分支走迂回的路径，其长度使信号延迟的时间正好等于双脉冲两次脉冲间隔的时间。然后两根轴突又重新分叉，这两根轴突各有一根分支终结于一个抑制神经元上，但是从延迟分支来的叉分裂成直接和迂回的路径各一条。在单脉冲输入时，这个神经元的输出是没有，而在双脉冲输入时，（在延迟的位置上）就有一个双脉冲。携带这一输出的轴突分叉成3个分支，所有分支都以抑制突触结终止于最后一个兴奋神经元上。原先分叉的轴突余下的两部分，每一根再分成两根，所有4根分支以兴奋突触结终止于最后这个神经元上。读者也许愿意去检查，这最后的兴奋神经元是否提供了所必需的“非”输出（也就是如果输入是单脉冲则输出为双脉冲，输入若是双脉冲则输出为单脉冲）。（这一方案仿佛过于复杂，但是我已经尽力而为了！）读者可以自己消遣，为上面其他逻辑门提供直接的“神经元”构造。

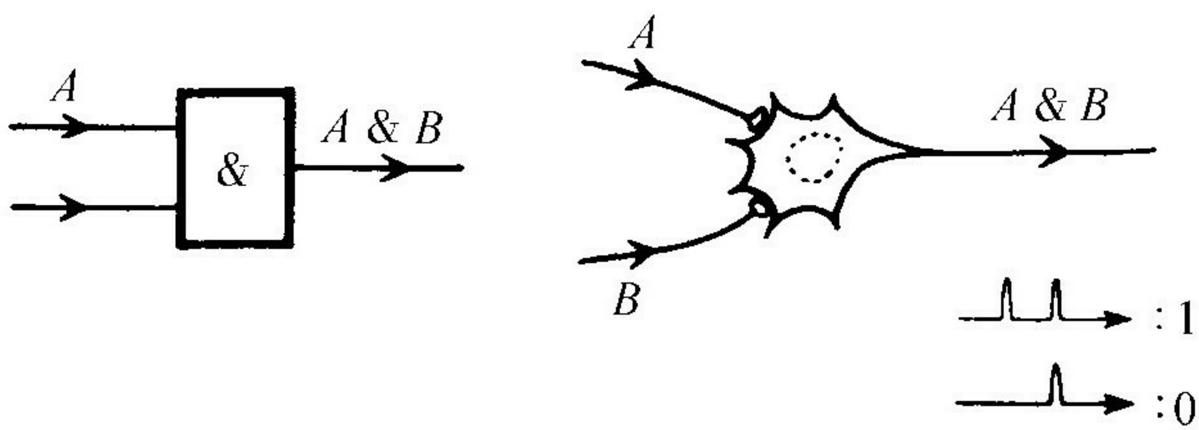


图9.13　一个“与”门。在右边的“神经元模型”中，只有当输入达到单脉冲强度的2倍时，神经元才被激发

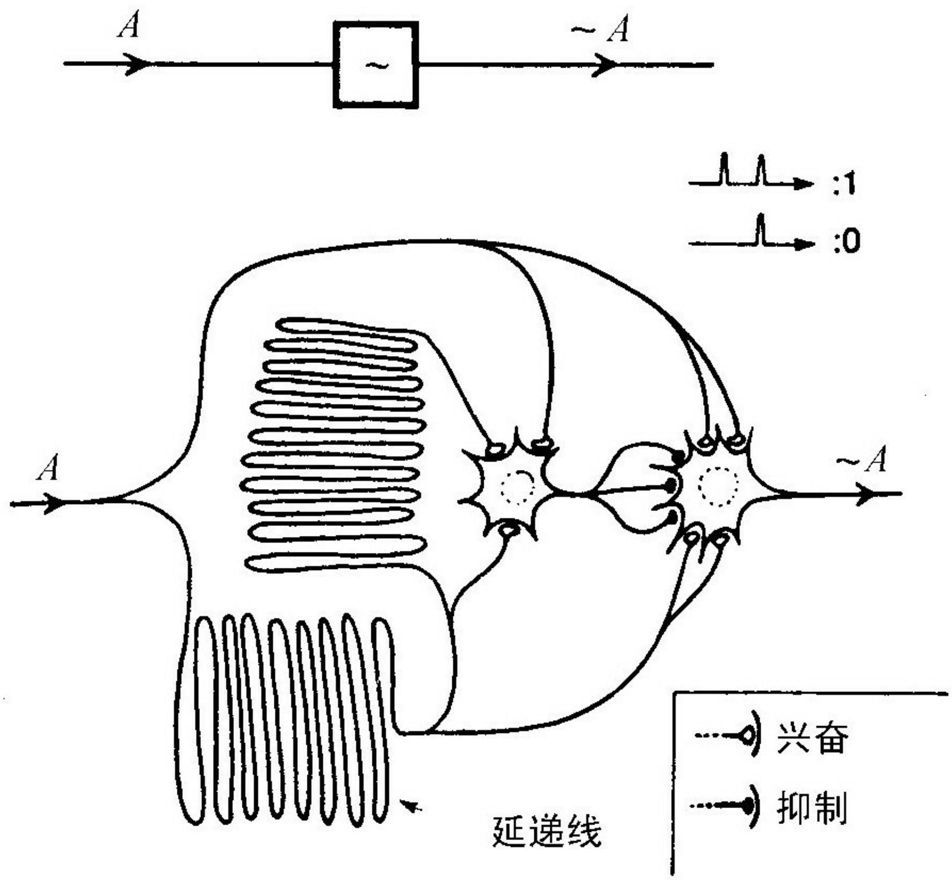


图9.14　一个“非”门。在此“神经元模型”中（至少）双倍强度的输入又是激发神经元所必需的

当然，这些显明的例子不能认真当作头脑真正详细行为的模型。我只想指出，我在上面给出的神经元激发模型和电子电脑构造之间具有本质上等价的逻辑。容易看出一台电脑可以模拟任何这类神经元相互联结的模型；而上述的详细构造表明以下这个事实，即神经系统反过来可以模拟一台电脑，并因此能像一台（普遍的）图灵机那样行为。虽然在第2章讨论图灵机时并没有用到“逻辑门”[7]，而事实上如果我们要去模拟一台一般的图灵机，需要比逻辑门更多的东西。假定我们允许自己用巨大却有限的神经元储库去近似一台图灵机的无限磁带，这样做并没有牵涉到新的原则问题。这似乎是在论断，头脑和电脑本质上是等价的！

但在我们过于轻率地下这个结论之前，应该考虑神经行为和现代电脑行为之间各种可能有意义的差异。首先，把激发神经描述成全有或全无的现象是有些过于简化。那现象是指沿着轴突移动的单脉冲。但事实上，当一个神经元“激发”时，它发射出一整串距离很近的脉冲。甚至在神经元不激发时，它也发射脉冲，只是以很慢的速率而已。当它激发

时，是连续脉冲的频率极大地提高。神经元激发还有随机的一面。同样的刺激不总产生同样的结果。此外，头脑行为并不需要电子电脑电流所需要的那么准确的计时；必须指出，神经元动作的最快速率为每秒一千次，比最快速的电子线路慢很多，大约慢10-6倍。还有，尽管我们现在知道头脑（在诞生时）联结的方式比50年前我们以为的方式更为精密，它不像电子电脑中非常准确的接线，神经元的实际连接仍有很多随机性和重复性。

上面的大部分内容好像是说，在比较头脑和电脑时头脑处于不利地位。但是头脑还有其他有利的因素。逻辑门只有很少的输入输出导线（譬如最多三四根），而神经元可以有大量的突触附在上面。（一个极端的例子，称为普奇涅细胞的小脑神经元大约有80000个兴奋突触末端。）还有，头脑中神经元的总数甚至超过最大电脑的晶体管总数——头脑中可能有10^{11} 个，而电脑中大约“只”有10^{9} 个！当然，电脑中的数目将来很可能增加 [8] 。而且，头脑细胞数目之大主要来自在小脑中发现的极大数目的小颗粒细胞——大约共有300亿（3×10^{10} ）个。如果我们相信仅仅是由于神经元的巨大数目就能使我们得到意识经验，而现代电脑不能具有意识，那么我们必须寻求更多理由来说明为何小脑行为显得完全无意识，而大脑跟意识有关，大脑神经元数目仅是小脑的2倍（大约为7×10^{10} ），其密度也更小得多。

头脑可塑性

在大脑行为和电脑行为之间还有其他不同之处。这就是与所谓头脑的可塑性有关。依我看来，其重要性超过迄今所提到的一切。认为大脑只是用导线联结起来固定的神经元组合，在实际上是不成立的。神经元之间的相互连接不像上述电脑模型中那样固定，它会随着时间不断改变。我不是说轴突或树突的位置会改变。大部分复杂的“接线”在人一出生就建立了大致的轮廓。我是指不同神经元的突触结实际发生联络的地方。这些经常发生在叫作树突棘的地方，那是树突上非常微小的突起，和突触结的接触可发生在这里（图9.15）。这里所谓“接触”不表示碰触，而是留下距离刚好大约1/40000毫米的狭缝（突触间隙）。现在按照一定的条件下，这些树突棘能缩小离开并且断开接触。或者它们（或新的）能长出并产生新的接触。这样，如果我们认为头脑中神经元连接实际上构成了一台电脑，那它就是一台能够一直随时变化的电脑！

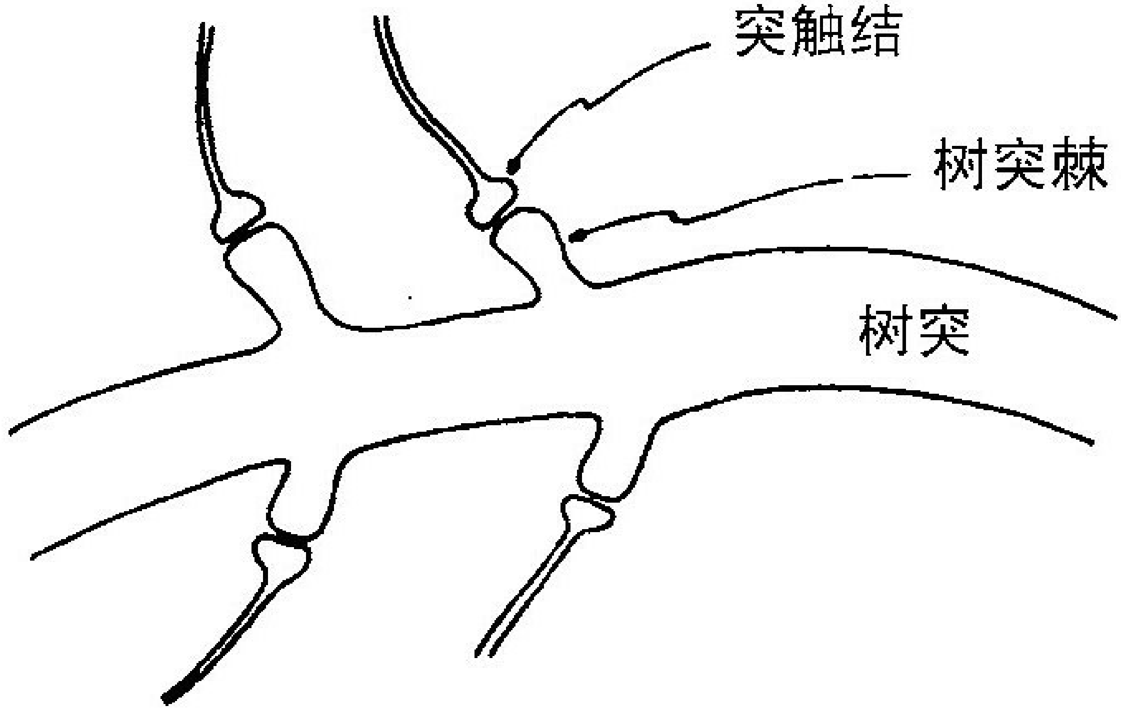


图9.15　突触结和树突棘。树突棘的长大和缩小很容易影响结的效果

根据长效记忆如何记录的主导理论之一，突触连接的这种变化正是提供储存必要信息的方法。如果真如此，那么我们就看到，头脑可塑性不仅是偶发的复杂性，而且是头脑活动的主要特征。

这些连续变化的基础机制是什么呢？这些变化可进行得多快？第二个问题的答案似乎很有争议，但是至少有一学派坚持可在几秒钟内进行这种变化。如果永久记忆的储存归功于这种变化，那是可以预料到的，因为记忆的记录确实是在几秒钟之内的事（参阅*Kandel 1976* ）。这对于我们下面的讨论有重大的含意。在下一章我们将回到这个重要的问题上来。

什么是头脑可塑性的基础机制呢？有一种天才理论（归功于*Donald Hebb 1954* ）提出，具有如下性质的某些突触（现在称作“赫伯突触”）：每当神经元A激发后跟着神经元B亦激发，在这两个神经元之间的赫伯突触就会被加强，否则就会被减弱。这和赫伯突触本身是否影响到B的激发没有关系。这引起了某种形式的“学习”。基于这种理论，出现了各种试图模拟一个学习/解决问题的活动的数学模型。这些被称为

神经网络。这类模型似乎的确具有某些基本的学习能力，但迄今它们离开头脑的实际模型还遥远得很。不管怎么说，控制突触联结变化的机制很可能比已经提及的机制更复杂，这很明显地需要我们去更深入地理解。

与突触结释放神经传递物相关的还有另一方面。这些释放有时根本不发生在突触裂，而是进入一般的细胞之间的液体，也许是为了影响非常远处的其他神经元。许多不同的神经化学物质似乎是以这种方式发射出来。而且有些记忆理论与我在前面指出的不同，这些理论依赖于不同种类可能涉及的化学物质。头脑状态肯定以一般方式受头脑其他部分产生而存在的化学物质（譬如荷尔蒙）的影响。神经化学的整个问题是复杂的，提供涵盖所有有关方面的可靠而精细的电脑模拟将是非常困难的。

并行电脑和意识的“一性”

许多人显然持这种意见，认为发展并行电脑是建立具有人脑功能的机器之关键。我们在下面简略地考虑这种目前流行的观念。并行电脑，与串行电脑相对比，能独立进行非常大数量分开的计算，而这些大体上独立运算的结果，断断续续地合并在一起，对整体计算作出贡献。建造这类型电脑的主要动机来自于模仿神经系统的运行，因为头脑不同部分的确似乎具有进行分开而独立计算的功能（例如，在视皮质处理视觉信息）。

在这里必须说明两点。首先，并行和串行电脑在原则上没有什么不同。事实上两者皆为图灵机（参阅第2章60页）。不同处只在于整个计算的效率或速度。有些类型的计算用并行组织的确更有效率，但并不总是这样。第二点，至少我自己的意见是，并行经典计算不可能掌握我们意识思维的关键。意识思维的一个显著特征是它的“一性’（至少当一个人处于正常心理状态，而且不是“头脑分裂”手术的患者时！），这和同时进行大量独立活动成显明对比。

类似“你怎么能期望我同时想两件事情呢？”的抱怨乃是司空见惯的事。一个人的意识究竟能同时进行许多不同的思考吗？也许有人能同时进行一些思考，但是与其说像是同时、有意识地、独立地实际思考不同

的题目，不如说像是在这些题目之间跳来跳去。如果一个人在意识中完全独立地想两件事，甚至哪怕是在短暂的时间里，则就似乎具有两个分开的意识。而对于正常人而言，所能体验到的是一个单独的意识，该意识可以模糊地知悉许多事，但是在任一个时刻只能集中于一件特定的事情。

当然，我们这里所说的“一件事”的含义一点也不清楚。在下一章我们将在庞加莱和莫扎特灵感中遇到一些“单独思想”的非常显著的例子。但是，为了辨识一个人在任何时刻都可能有非常复杂的意识，我们不必舍近求远。例如，想象一个人决定晚饭要吃什么。这样一个意识思维之中牵涉到大量信息，而且要用相当长的言语才能完全描述清楚。

对我来说，这种意识认知的“一性”似乎和并行电脑图像相去甚远。另一方面，那个图像也许更合适作为头脑无意识行为的模型。不同的独立动作——散步、扣纽扣、呼吸或者甚至讲话可同时多多少少自动地进行，人们不必在意识上感觉到任何动作在进行！

另一方面，我认为在意识的“一性”和量子平行主义之间可以想见具有某种关系。我们记得量子理论中，在量子程度上允许不同选择在线性叠加中共存！这样，一个单独的量子态在原则上可由大量不同的，而且同时发生的活动组成。这就是所谓的量子平行主义。我们很快就要考虑“量子电脑”的理论观念，这样的量子平行主义在原则上可用于同时进行大量的计算。如果意识的“心理状态”在某种形式上和量子态同类，那么思维中某种形式的“一性”或整体性对量子电脑就比对普通并行电脑更为适合。这个观念中有一些方面引人注意，我在下一章再回到这上面来。但是在认真接纳这个思想之前，我们必须提出以下问题，就是量子效应究竟和头脑活动有何相关。

量子力学在头脑活动中有作用吗

上面有关神经活动的讨论全部都是经典的。除了迄今必须提起的一些物理现象，其基础的机制必须包含一部分量子力学的因素以外（例如离子，以及它们的单位电荷、钠和钾门、决定神经信号开关特性确定的化学势、神经传导物的化学作用）。真正量子力学控制在某些关键处还有更清楚的作用吗？如果上一章结尾的讨论不是无的放矢的话，结论似

乎是肯定的。

事实上，至少在一个明显的地方，单量子水平的作用对于神经活动很重要，这就是视网膜。（我们记得视网膜事实上是头脑的一部分！）以蟾蜍做的实验显示，在适当条件下，一颗单独光子打到已适应黑暗的视网膜上就足以触发一个宏观的神经信号（*Baylor, Lamb, and Yau 1979*）。这也适用于人眼（*Hecht, Shlaer, and Pirenne 1941*），但是在此情况下还存在额外的压抑这种弱信号的机制，使得它们不会由于太多的视觉"噪声"而混淆了感觉到的视像。为了能使已适应黑暗的人实际上得知光子的来临，大约需要7颗光子的组合的信号。尽管如此，对单光子敏感的细胞的确存在于人类的视网膜中。

既然在人体中存在单量子的情况下就能触发的神经元，寻找人脑主要部分何处能发现这类细胞就是很合理的了，据我所知，对此还未找到证据。所有考察过的细胞类型都有一个临界值，要激发该细胞就得需要大量的量子。然而人们猜测，在头脑的某一深处可望找到对单量子灵敏的细胞。如果证明情形的确如此，则量子力学对头脑活动的意义就非常重大。

即便如此，在这里量子力学还不显得非常有用。这是由于量子只是一种用来激发信号的手段。没有得到量子特有的干涉效应。我们从这些得到的，最多似乎只是确定一个神经元是否会激发，这很难看出对我们有多大用处。

然而，这里牵涉到的问题不是那么简单。让我们重新考虑视网膜。假定从一半镀银的一面镜子反射来一颗光子到达视网膜。它的状态涉及以下状况的复线性叠加：光子打到视网膜细胞和光子没打到视网膜细胞，譬如穿过窗户飞到空中去（参阅图6.17，第323页）。到达它可以打到视网膜的时刻，只要量子理论的线性规则U（也就是薛定谔态矢量演化，参阅318页）成立，则我们就能得到有神经信号和没有神经信号的复线性叠加。当它作用到主体的意识上时，两个不同选择中只有一个被感知发生，这时另一个量子步骤R（态矢量减缩，参见318页）应该起了作用。（我在此不理会多世界观点，它本身有许多问题！参阅378页。）连同上一章结尾触及的考虑，我们应该问，信号的通过是否扰动了足够的物质，达到那一章的单引力子水平。虽然把光子能量转变成实在信号中的物质运动时，视网膜的放大效应真是令人印象深刻——运动

质量的放大也许达到10^{20} 倍——但这个质量仍比普朗克质量m_p 小许多数量级（譬如大约为10^8 ）。然而，一个神经信号在它周围产生了可以探测得到变化的电场（一个以神经为轴，沿着神经运动的圆环形电场）。这场会显著地扰动周围环境，单引力子标准在这些环境中可容易地达到。这样，按照我提出的观点，R过程在我们感知或没有感知闪光之前早就已经进行过了。由此观之，我们的意识对于态矢量减缩不是必要的！

量子电脑

如果我们猜测在头脑深处对单量子敏感的神经元会有重要的作用，我们就想知道他们会有什么效应。首先我将讨论德义奇的量子电脑概念（参阅第4章191页），然后看看是否和这里的讨论有相关之处。

正如前面指出的，其基本概念是利用量子平行主义。根据这个原理，两个完全不同的事情应当被认为在量子线性叠加中同时发生。正如光子被半面镀银的镜子反射，同时光子又穿过镜子或者是通过两个缝隙中的每一个。对于量子电脑，这两个叠加的不同情况就是两个不同的计算。我们对两个计算的答案不感兴趣，而是对利用从这对叠加抽取出的部分资料感兴趣。最后，当两个计算都完成时，对这些计算进行适当的“观察”以得到必需的答案[9] 。仪器用这种同时进行两个计算的办法来节约时间！迄今这个方法并没获得什么重大好处，这是因为可以想见利用一对分开并行的经典电脑（或一台单独的经典并行电脑）比用量子电脑更直截了当得多。然而，量子电脑可能要到需要非常大量的（也许是无限大的数目）并行计算时才会有真正的好处。我们对个别计算的答案不感兴趣，而对所有答案适当的组合感兴趣。

量子电脑的建造在细节上会涉及量子形式的逻辑门，其输出为应用在输入上某个“幺正运算”的结果。这是U作用的一种情形，而电脑所有的运行就是u过程进行到最后阶段，直到最后的“观察行为”R为止。

根据德义奇的分析，量子电脑不能用来进行非算法的运算（也就是超越图灵机功能的事）。但是在非常巧妙的设计情形下，在复杂性理论意义来说（参阅184页），它能比标准的图灵机获得更大的速度。对于这么杰出的设想，目前的结果仍有点令人失望，但是这有点言之过早。

量子电脑和包含极多数目单量子敏感神经元的头脑，两者行为会有什么关系呢？这个类比的主要问题是量子效应在“噪声”中很快就消失了——头脑太“热”不能在足够长的时间内维持量子相干性（就是通常可以用U的连续作用有效地描述的行为）。以我的术语，这表明连续达到单引力子的水平，使得R作用持续不断地进行，其间穿插着U演化。

我们期望量子力学对了解头脑有所帮助，但目前看来希望并不大。也许我们注定只是电脑而已！我个人不这么认为。但是如果我们要找到答案，就必须要更深入地思考。

超越量子理论

我希望能回到这本书的基本论题上来。我们的世界图像是由经典和量子理论的法则所制约的，就现在所理解的这些法则而言，这图像足以描述头脑和精神吗？对我们头脑的“通常”的量子描述一定存在一种困惑，因为“观察”行为被当成解释传统量子理论的要素。是不是只要思维或知觉一旦进入意识，“头脑”就被认为在“自我观察”？量子力学如何顾及这一点并应用到整个头脑，传统理论没有提供我们明确的法则。我曾试图为R作用提供一个和意识完全无关的判据（“单引力子水平”），如果类似的判据能发展成完全连贯的理论，那就出现一个用量子描述头脑的方法，比迄今存在的描述都更清楚明白。

然而，我相信不仅在我们试图描述头脑行为时才引起这些基本问题，数字电脑本身的作用必需依赖量子效应。依我的看法，这些效应并没有完全摆脱量子理论根本的困难。这种“必需的”量子依赖性是什么呢？为了理解量子力学在数字电脑中的作用，我们首先必须问，如何使完全经典的物体能像数字电脑那样行为。我们在第5章考虑了弗列得钦——托弗里经典的“台球电脑”（参阅221页）；但是我们还注意到这个理论“仪器”有赖于某种理想化，这种理想化回避了经典系统中固有的不稳定性问题。这个不稳定性问题可在相空间中被描述成随着时间演化的弥散（第235页图5.14），导致经典仪器运作的准确性几乎不可避免地连续损失。能终止这种准确性降级的最终是量子力学，现代电子电脑分立态的存在是必需的（譬如以数字0和1来编码），这使得电脑处于此态或彼态一清二楚。这是电脑操作的“数字”性质的要素。这种分立性最终有赖于量子力学。（我们还记得能级、谱频率、自旋等的量子分立性，

参阅第6章。）甚至老的机械计算仪器也依赖于不同零件的坚固性，而坚固性实际上也有赖于量子理论的分立性[10]。

但是，不仅从U的作用才能得到量子的分立性。其实，薛定谔方程在防止不想要的弥散和“精度损失”方面比经典物理方程更糟！根据U的时间演化，一个单粒子原先空间定位的波函数会散开到越来越广的范围去（参阅320页）。如果不是R的作用时时发生的话，更复杂系统有时也遭受到这种不合情理的无定域性（回忆薛定谔猫的例子！）。（例如，原子之分立态具有确定的能量、动量和总角动量。一般“散开”的态是这种分立态的叠加。正是R的作用在某阶段使原子实际“成为”这些分立态之一。）

我认为，经典力学不能解释我们思考的方式。如果没有一些根本改变使R成为“实在”过程，连量子力学也不能解释。也许连电脑的数字行为都需要对U和R之间相互关系有更深入的理解。至少我们知道电脑（由于我们的设计！）的行为是算法的，而且我们不想采用任何物理定律中推想出来的非算法行为。但是，我坚持头脑及思维的情形是非常不同的。在（意识）思考过程中包含非算法的要素是说得通的。我在下一章将探讨我相信有这种要素的理由，以及猜测究竟是什么了不起的物理效应构成影响头脑行为的“意识”。

第10章　精神物理的寻求

精神是做什么的

在讨论精神——身体问题时，通常有两个不同的问题受到关注：“物质物体（头脑）实际上如何引发意识？”以及相反的命题，“意识的意志行为在实际上如何影响（显然由体力驱动的）物质物体的动作？”这些是精神——身体问题被动和主动两方向。我们的“头脑”（毋宁讲“意识”）中显然有种非物质的“东西”，一方面它受物质世界召唤，另一方面又能影响物质世界。然而在这最后一章的初步讨论中，我宁愿考虑有点不同或许却更科学的问题，它和主动及被动的问题都有关系，希冀我们的探索能进一步理解这些根本的古老哲学难题。我的问题是：“意识赋予实际拥有它的人们哪些选择性上的好处？”

在以这种方式表达该问题时，涉及几个隐含的假设。首先，有人相信意识实际上是一种可以科学描述的“东西”。这里假定这个“东西”实际上“会做一些事情”，而且其作为对拥有它的生物有助，所以其他没有意识的同等生物行为就不是那么有效。另一方面，人们也许相信，意识只不过是在足够复杂的控制系统中被动的伴随，而它自身实际上并无任何作为。（例如，强人工智能支持者就采取这种观点。）另外的看法是，在意识现象中或许存在某种神圣或神秘的目的，可能是还未被我们揭示的目的。仅仅按照自然选择的思想去讨论这个现象会完全忽视这个“目的”。我的思考方式有点倾向于这种论证的更科学的形式，即是所谓的人存原理。该原理断言，我们存在其中的宇宙的性质受到如下强烈的限制，即必须存在像我们这样有知觉的生物以便对它进行观察。（在第8章448页已经稍微提到了这个原理，下面还要进一步讨论。）

我将依序讨论这些问题的大部分。但是，我们首先应该注意到，“精神”这个术语在我们提及“精神——身体”问题时也许有点误导。人们经常讲到“无意识的头脑（精神）”。这表明我们认为“精神”和“意识”两个术语不是同义词。也许当我们提到无意识的精神时，我们有个模糊印象认为“后面有人”在幕后活动，但他通常不直接触及我们的感知（也许除在睡梦、幻觉、着迷或弗洛伊德事故以外）。也许无意识的精

神实际上自己有知觉，但是在正常情况下这知觉和平常我们所指头脑中“我”的那部分完全分离。

这也许不像初看起来那么强词夺理。某些实验指出，甚至病人在全麻醉状态下被动手术时，还存在某种“知觉”。例如，当时进行的谈话会在以后“无意识地”影响病人，以后在催眠下有时能回忆起这些谈话，如同当时实际“体验到”似的。此外，被催眠暗示阻挡于意识外的知觉在进一步催眠之后可以当作好像“体验过的”被回忆起来，但似乎是“处于不同的意识轨道上”（参阅*Oakley and Eames 1985* ）。尽管我猜想，赋予无意识精神任何通常的“知觉”是不正确的，我对这些问题一点也不清楚，而且我也不想在这里讨论这些猜测。尽管如此，在意识和无意识的精神之间进行划分，肯定是一个既微妙又复杂的问题，我们以后还要涉及。

让我们尽可能直截了当地讨论，我们所指的“意识”是什么以及我们相信它在什么场合存在。在目前理解的程度上，我认为试图为意识提供一个准确的定义是不明智的。但是，我们可以充分仰仗我们主观印象和直觉常识来解释这个术语的含义，以及何时这种意识的特质会呈现出来。当我自己处于意识状态时，我或多或少是知道的，而且我认为其他人也有相同的经验。我处于有意识时，我似乎必须意识到某种东西，也许是感觉，诸如痛、温暖或者彩色风景、音乐之声；或者我意识到诸如迷惑、沮丧或快乐的感情；或者我可以意识到某些过去经验的回忆；或者理解其他人讲什么或是自己的一个新思想；或者我意识地想发言或采取行动如从座位上站起来。我还可以“后退一步”意识到这些企图，或者自己痛的感觉，或者自己记忆的经验，或者自己获取的理解，或者甚至只是对自己意识的意识。假如我正在做梦的话，在睡觉中也可以具有某种程度的意识；或许当我快醒过来时，我有意识地影响那个梦的发展方向。我预备相信，意识只是程度上的差别，而不是全部无或全部有。我把“意识”这个词和“知觉”基本上当成同义词（虽然“知觉”也许只比我所指的“意识”被动一点），而“精神”和“灵魂”则有更多内涵。目前这两词的定义更加不清楚一些。我们在解释“意识”时就够麻烦的，所以如果我不触及“精神”和“灵魂”更深入的问题的话，希望读者能够原谅！

还有一个问题就是“智慧”这个术语表明什么。毕竟人工智能专家要关心的是智慧，而不是更模糊的“意识”的问题。阿伦·图灵（1950）在其著名的论文中（参阅第1章第5页）没有这么直接地提到“意识”，但是

提到了“思维”，并且在标题上用“智慧”这个词。依我自己看待事物的方法，智慧的问题属于意识的问题范围内。我相信，如果没有意识相伴随，真正的智慧是不会呈现的。另一方面，如果人工智能专家最终能模拟不存在意识的智慧，则在定义术语“智慧”时应该包括这种模拟智慧才会令人满意。在这情形下我真正关心的不是“智慧”问题。我首先要关心的是“意识”。

当我断言自己相信真正的智慧需要意识时（由于我不相信强人工智能的只要制定一个算法即能召唤起意识的论点），根据我们现在术语的意义，我的意见暗示智慧不能用算法的方法，也就是电脑，正确地模拟智慧。（参见第1章关于图灵机的讨论。）因为我很快就要有力地论证（特别是参看第523页以下三节有关数学思维的讨论），在意识行为中必须有本质的非算法成分。

下面让我们讨论，某种有意识的东西和另一种在其他方面都“等效”而无意识的东西是否有操作上的差异。某些意识的物体永远会呈现它的存在吗？我想，对这个问题必须回答“是”。然而，因为对动物王国中何者有意识完全缺乏共识，所以我的信念几乎得不到任何赞同。有些人根本不允许非人类动物拥有意识（还有人甚至不允许早于公元前1000年左右的人类拥有意识，参阅Jaynes 1980），而另外的人则赋予昆虫、蛆虫甚至砖头意识！至于我自己则怀疑蛆虫或昆虫会有，而岩石肯定不会具备这种品质。但是，一般来讲，我觉得哺乳动物真的具有一些知觉。我们由缺乏共识至少可以推论，没有一般可以接受的呈现意识的标准。不过仍然可能存在一种意识行为的标志，只是还没被普遍承认而已。尽管如此，这也只是标明意识的主动作用。若没有相关活动的对象，很难看出何以直接确定知觉的存在。这是从以下悲惨的事实中得知的。在20世纪40年代，对年幼孩子动手术时用箭毒来进行“麻醉”，而箭毒的实际效用是麻痹肌肉上的运动神经，所以这些不幸的孩子实际经历了在当时外科医生不可能知道的灾难（参阅*Dennett 1978*， P209）。

让我们转向意识可能具有的主动作用。意识有时候的确能够具有操作上辨别得出的主动作用，但必须如此吗？我相信这种说法的原因有点与众不同。首先，利用我们的“常识”，我们经常觉得能直接知觉到他人实际上有意识。那个印象不太可能会错[67]。虽然有时候一个有意识的人（正如受了箭毒的小孩）看起来不明显，但是一个无意识的人更不可能显得有意识！所以必须有一种行为模式作为意识的特征（尽管不一定

被意识所证实），而我们可以通过自己的“常识直觉”可以敏锐地感觉到它。

第二点，考虑到自然选择的无情过程。正如我们在上一章所看到的，意识不能通达至头脑所有的活动。的确，“较老的”小脑中神经元的局部密度极大，小脑似乎进行着意识根本不直接参与的非常复杂的行为。然而大自然已经选择演化像我们这样有知觉的生物，而不愿演化利用完全无意识的控制机制来指示行为的生物。如果意识没有选择的目的，而像小脑这样没有知觉的“自动”头脑似乎也能达到目的时，大自然为何要这么不厌其烦地去演化意识的头脑呢？

此外，有一个简单的“底线”原因令人相信意识必须具备某种主动效应，即使这效应不是一种选择优势。否则为什么像我们这样的生物有时候，尤其是在探索此事时，会被“自我”的问题所烦恼呢？（我可以说：“你为何在读这一章？”或者“为什么起先我会有强烈欲望要写这本论题著作”？）很难想象一台毫无意识的自动机会为这类想法而浪费时间。另一方面，由于意识生物似乎有时以这种滑稽的方式行为，因此他们的行为和如果他们没有意识时不同。这样，意识具有某种主动效应！当然，特意给电脑编一道程序使之显得以这种可笑的方式行为是很容易的事（例如，程序指使它到处乱走并咕哝：“啊，亲爱的，什么是生活的意义？为何我在这里？我所感到的这个‘自我’究竟是什么？”）。显然地，当无情的丛林自由竞争早就应该根除这种无用的废物，可是为何物竞天择却偏爱人类这种生物！

对我来说，有一点是很清楚的，即当我们（或许是暂时的）作为哲学家时所热衷的沉思和喃喃自语本身并非由选择而来，而是确实有意识的生物必须背负的“包袱”（从自然选择的观点），并且这生物的意识是经过自然选择而来，不过是由于其他不同而又非常有力的理由。这包袱不太有害，而且很容易背负（即使不很情愿），我猜是因为自然选择不屈不挠的力量所驱使的。也许因为幸运的人类时而享有和平与繁荣，使我们不必总是为求生存而与自然环境（或与邻居）作战，我们才能开始对包袱内的宝藏神迷目眩。一个人正是看到他人用这种奇怪的哲学方式行事，才得以信服他是和除了自己以外确实具有精神的个体打交道。

意识究竟是做什么的

让我们接受这样的观点，即在生物中意识的存在实际上是使该生物具有某种选择优势。其特别优势会是什么呢？我曾听到这样的观点，一个掠捕者把自己当成猎物以猜想它下一步最有可能做什么，对掠捕者而言，知觉是一种优势。把自己想象成为该猎物，就能得到优势胜过它。

这种思想中很可能有一部分真理，但是我对此很难苟同。首先是假定猎物本身方面具有某种预先存在的意识，这样又把自己想象成一台“自动机”根本没有帮助，既然一台自动机按照定义是无意识的，不可能是“活”的东西！无论如何，我可以同样容易想象，一个完全无意识的自动掠捕者可以把它的自动猎物的实际程序作为子程序包含在自身的程序之中。我觉得，把意识牵涉到这种掠捕者——猎物的相互关系中根本没有逻辑上的必要。

当然，很难了解自然选择的随机过程怎么会聪明得将这猎物的程序的完整的复本给予自动机掠捕者。这听起来与其说是自然选择不如说是间谍活动！而部分程序（图灵机的一段磁带或某种与图灵机磁带近似的东西）对于一个掠捕者没有多大的选择优势。拥有整盘磁带或至少拥有整个自足的部分磁带是不太可能的。所以另一种可能性是，以下的观念或许具有一些真理，也就是从掠捕者——猎物这思路可以推论出某些意识的因素，而不仅是一个电脑方程式。但是，这里并没有抓住意识行为和“程序”行为之间实际的差别是什么的要点。

上面提到的观念和人们经常听到的意识观点相关，也就是一个系统如果本身具有某种东西的模型时才会“知觉”到该东西，而当它本身具有它自己的模型时才能“自我知觉”。但是，在一段电脑程序中包含另一段电脑程序的描述（譬如一段子程序）并没赋予第一段程序对第二段的知觉。电脑程序的自我参照也不会导致自我知觉。尽管经常听到这种断言，我的看法是，这类讨论尚未能触及知觉和自我知觉的真正问题。一台录像机对之所录下的风景没有知觉；对着镜子的录像机也不具备自我知觉（图10.1）。

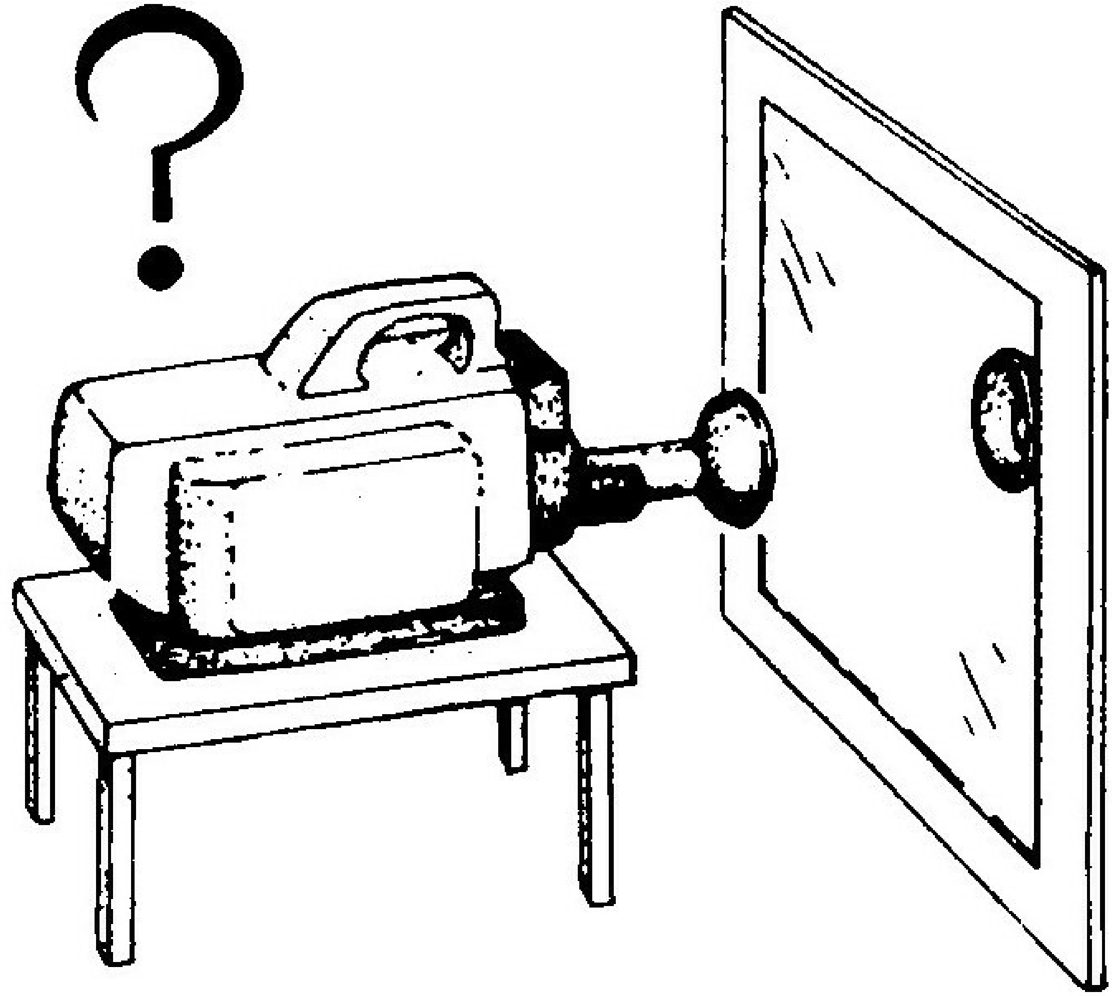


图10.1　对着镜子的录像机在自身中形成自身的模型。这使它具备自我知觉吗？

我想沿着不同的思路进展，我们已经看到，我们头脑中进行的活动不是全部伴随着意识知觉的（尤其小脑的活动似乎是没有意识的）。我们意识思维能做而无意识状态所不能做的事情是什么呢？这问题由于以下事实而变得更加无从捉摸，任何原先需要意识的事显然都可学得会并能无意识地（也许由小脑）执行。可以这么讲，当我们必须形成新的判断以及当预先还没形成习惯时，意识是必需的。想很精确地区别何种精神活动需要意识不是很困难的。也许，正如强人工智能的支持者（以及其他人）所坚持的，我们在“形成新判断”时是在应用某些定义得很好的，却难以了解的“高层次的”算法规则，而我们还未能知道其运作的方式。然而，我想认为有些术语，用来区别有意识和无意识精神活动的，至少可作区分非算法和算法的参考：

需要意识的	不需要意识的
“常识”	“自动的”
“真理的判断”	“盲目地跟随规则”
“理解”	“编程序的”
“艺术鉴定”	“算法的”。

这些区别也许不那么一清二楚，尤其是因为许多无意识的因素进入我们意识的判断之内：经验、直觉、偏见甚至我们逻辑的正常运用。但是我要宣称，判断本身是意识行为的呈现。所以我提出，头脑的无意识行为是按照算法过程进展，而意识的行为则完全不同，它以一种不能被任何算法所描述的方式进展。

颇具讽刺意味的是，我在这里提出的观点和我经常听到的其他观点几乎刚好相反。人们经常说，有意识的头脑以一种“理性的”，人们可以理解的方式行为，而无意识是神秘的。从事人工智能的人们经常宣称，只要能理解意识思维的某些方法，人们就能知道如何让电脑照做，而人们对神秘的无意识过程尚没有解决之道。按照我自己的推理，无意识过程很可能是算法的，但是该算法是极其复杂的，要仔细解开它极为困难。完全有意识的可以合理解释为完全合逻辑的思维，也可以（经常）用算法来表达，但是在完全不同的水平上。我们现在不去思考内部的功能（如神经元激发等）而是整个思想的运作。这种思想运作有时具有算法的特性（正如早期的逻辑：由亚里士多德所表达的古代希腊演绎法或者是数学家乔治·布尔的符号逻辑；参阅Gardner 1958），有时它不具有这类特征（正如哥德尔定理以及在第4章所举的例子）。我现在宣称的判断的形成是意识的标志，人工智能专家不知如何用电脑为它编出程序来。

人们有时反对，说这些判断的判据毕竟不是有意识的，为什么我要认为这类判断起因于意识呢？但是，这样问就错过了我想要表达的思想要点。我不是要我们有意识地理解我们如何形成意识印象和判断，这会导致我刚提到那种水平上的混淆。构成我们意识印象的原因是意识无法直接触及的。这些必须用比我们现知的实际思想更深的物理水平来考虑。（我要在下面提出设想！）意识印象本身就是（非算法的）判断。

我们的意识思维应该有非算法的性质，这是我们前面章节的基本主题。尤其是从第4章有关哥德尔定理的论证得出的结论指出，意识的沉思至少在数学方面有时能够使人用算法不能做到的方式去确定某个陈述的真理性。（我就要仔细地阐释这个论证。）算法本身的确从未确定真理！要使算法只产生谬误和只产生真理一样容易。为了确定一个算法有效与否，人们需要有外在的洞察（后面还要讲到）。我的论断是，在适当的环境下从错误中判断出真理的能力（或从丑恶中得到美丽），正是意识的标志。

然而，我要讲清楚，我不是指用魔术式的“猜测”。意识对于猜乐透号码（公平进行的）毫无用处！我是指在人们处于意识状态下连续进行的判断。把所有相关的事实、感觉印象、记住的经验都集中在一起，把事物相互衡量，甚至有时形成灵感的判断。原则上，只要得到足够的信息即能作有关的判断，但是由混乱的数据中抽取需要的并形成适当的判断，也许没有清楚的算法的步骤存在，或者虽然存在，但不切实际。也许我们会有这种情形，一旦做出判断，去检查该判断是否准确比当初形成该判断更像算法的过程（或许只是一种更容易的过程）。我猜想在这种情况下，意识本身会成为召来适当判断的方法。

为什么我说判断的非算法形成是意识的标志呢？有部分原因来自于我自己作为数学家的经验。当我的知觉未能充分注意无意识的算法行为时，我就是不信任它们。在进行某些计算时，把算法当作算法通常是没错，但是对研究中的问题这算法是正确的选择吗？一个简单的例子是，人们学会把两个数乘在一起以及把一个数除以另一个数的算术规则（或人们宁愿借助于算法袖珍计算机），但是人们面对这类问题时何以得知应该乘或除这些数呢？为此，人们需要思考并做出意识判断。（我们很快就会看到，为什么这样的判断至少有时候必须是非算法的！）当然，一旦人们做过大量类似问题，这些数该乘还是该除会变成第二天性，而且可以算法地执行——也许由小脑。在那个阶段不再需要知觉，而且足

以放心让他的意识精神去琢磨或沉思其他事情——不过人们必须不时检查该算法是否被导入（即使很细微的）歧途。

相同的情形在所有水平的数学思维过程中不断发生。当人们在进行数学过程时通常竭力寻找算法，但是这种努力的过程本身并不是一种算法过程。在某种意义上，一旦找到一个合适的算法，该问题就解决了。此外，用数学来判断某些算法是否精密或合适，需要很多的意识的关注。第4章内描述数学形式系统的讨论也提过类似的情形。人们可从一些公理开始推论出各式各样的数学命题。后者的步骤可以完全是算法的；但是需要一位有意识的数学家去判断这些公理是否合适。在往后第二节讨论中可清楚得知，这些判断必须不是算法的。但在此之前，让我们考虑一种更盛行的有关我们头脑功能及其起因的观点。

算法的自然选择

如果我们假设人类头脑的行为，不管是意识的还是无意识的，只是在执行一种非常复杂的算法，那么我们应该询问这种非常有效的算法从何而来。标准答案当然是“自然选择”。具有头脑的生物在进化时，那些算法比较有效的会有更好的生存倾向，并因此在总体上有更多的后代。由于这些后代从其父母遗传到比较好的算法，因此也比其堂表亲戚带有更有效的算法，所以算法就这样被逐渐改进。由于生物在进化时可能有很多断断续续的现象，所以这种改进不必是稳定的。生物进化可以到达了不起的阶段，就像我们可以（显然地）从人类头脑看到的（比较*Dawkins 1986*）。

甚至根据我自己的观点，由于我想象头脑许多行为的确是算法的，正如读者从上述讨论中所推论出来的，这个图像必须包含某些真理。而且我强烈相信自然选择的威力。但是我看不出自然选择本身如何能演化算法，这种算法能有意识地判断我们似乎拥有的其他算法是否有效。

想象一道平常的电脑程序。它怎么会变成活的呢？显然不能（直接地）由自然选择而来！要有电脑程序人员构思写出这程序并确认它会执行所预定的步骤。（实际上，大多数复杂的电脑程序都含有错误，通常很小，但常常微妙得除了在非常的情形下不会被发现。这种错误的存在不会严重改变我的论证。）有时一个电脑程序本身可由另一个程序（譬

如由“主导”电脑程序）“写出”，但是那个主导程序本身是人类才智和洞察的产物；或者该程序只是从许多其他电脑程序的产物拼凑而成。但是在所有的情形下，程序的有效性和概念本身最终要归功于（至少）一个人类的意识。

当然，人们可以想象情况也许不必如此，只要有足够的时间，电脑程序可能会自动由某种自然选择的过程演化得来。如果人们相信，电脑程序人员的意识行为本身就是算法，那么他实际上应该相信，算法已用这种方法演化至今。然而使我忧虑的是，一个算法有效性的决定本身不是一个算法过程，我们在第2章已经看过这情形。（一台图灵机实际上会不会停，这问题不是用算法能够决定的。）为了决定一个算法实际上行不行，人们需要的是洞察，而不是另一个算法。

尽管如此，人们仍能想象某种自然选择过程可以有效产生近似有效的算法。然而，我个人很难相信这种可能。任何这类选择过程只能作用于算法的输出[68]，而不直接作用于算法行为的基础思想上。这不仅极无效率，而且我相信这肯定行不通。首先，仅从考察其输出是很难确定一个算法究竟是什么。（要构造两个完全不同的简单的图灵机行为非常容易，使两者的输出磁带到第2 65536位都是一样的，这个差异在整个宇宙的历史中永远不会被觉察出来！）此外，一个算法最微小“变化”（譬如一台图灵机在规格上或在它的输入带上轻微的改变）就会使之变成完全无用，很难看出随机的方式如何能产生算法实际的改善。（如果不知道其“意思”的话，甚至故意的改善也是困难的。这一点尤其被如下时常发生的情形所证实。当一道没有说明清楚或者复杂的电脑程序要作一点改变或改正时，而原先的程序人员刚好离开或死去，人们与其试图解开在该程序中暗含的意义和企图，不如将其丢弃后重写可能还容易些！）

也许可以设计更“健全”的方法来详细说明算法，使它避免上述的批评。在某些方面，这正是我自己要说的。这种“健全”的说明是算法的基础观念。但是观念，就我们所知，是需要意识精神来表明的东西。我们回到了意识究竟是什么的问题，什么是它能做而无意识的主体不能做的，自然选择究竟如何聪明，以至于得进化出那个最特异的品质。

自然选择的产物确实惊人。有关人类和其他生物的脑如何作用，我得到的一点认知使我充满了无以言喻的惊奇和赞美之情。单独神经元的

作用是非同寻常的，在我们出生之时为了以后需要担负的任务，这些神经元本身以惊人的方式以极大数量的连接组织在一起。不仅是意识本身，而且必须用来支持意识的配件也都是那么令人印象深刻！

如果我们能发现，究竟什么品质可以使一个实体成为有意识的，那么我们就可能为自己建造这样的物体。虽然它们可能不符合我们现在所谓的“机器”之词意。因为这些物体是为了我们目前的任务，也就是为了获得意识而特别设计，所以可想而知，它们比我们更优越得多，它们不必从单独的细胞长大。它们也不必负担它们祖先的“包袱”（头脑或身体内一些老和“无用”的部分，还在我们身上留存全是因为我们远祖演化的“事故”）。人们可以想象，从这些优点看来，这类物体可以获得超越人类的成功，（依我等意见）算法电脑注定只能屈于卑微的地位。

但是，还有更多关于意识的问题。我们的意识也许某方面的确有赖我们的遗传和几十亿年下来的实际进化。我的想法是，进化明显具有“探求”未来的目的，所以它仍有神秘之处，事情至少看起来组织得比仅仅基于瞎碰机会的演化和自然选择要更好。也有可能这种表象完全是骗人的。物理定律作用的方式似乎有种因素使得自然选择的过程比单凭任意定律的过程更有效得多。其导致的“智慧探求”是一个有趣的问题，我将很快回到这问题上来。

数学洞察的非算法性质

正如我早先陈述过的，令人相信意识能够非算法地影响真理判断的大半原因是通过考察哥德尔定理而来的。如果在形成数学的判断时能看到意识的作用是非算法的，此时计算和严格证明构成非常重要的部分，则我们肯定会信服，在更一般（非数学）的情形下，这样非算法的因素对于意识也是关键的。

让我们回忆第4章用来建立哥德尔定理以及它与可计算性之间的关系的论证。这论证指出，不管数学家用什么（足够广泛的）算法去建立数学真理，或是类似真理的东西 [1] ，不管他采用什么形式系统去提供真理的判据，总有一些数学命题，譬如该系统显明的哥德尔命题P_k（k）（参考141页），这些算法不能提出答案。如果该数学家的头脑作用完全是算法的，那么实际用以形成他判断的算法（或形式系统）不能

用以应付从他个人算法建立起来的P_k （k）命题。尽管如此，我们（在原则上）能看到P_k （k）实际上是真的！既然他应该也能看得到这一点，这看来为他提供了一个矛盾。这个也许表明，该数学家根本不用任何算法。

这本质上就是鲁卡斯（1961）提出的论断，头脑的作用不能完全是算法的。但是时时有人提出许多相反的论点（例如，*Benacerraf 1967*；*Good 1969*；*Lewis 1969*，*1989*；*Hofstadter 1981*；*Bowie 1982* ）。我应该指出，在这里讨论的术语“算法”和“非算法的”是指一台普通电脑所能模拟的任何东西。这当然包括“并行运行”，还有“神经网络”（或是“连接机器”）、“启发”、“学习”（这里学习机器总是指预设好应该如何学习的固定步骤）以及和环境的相互作用（这可用图灵机的输入磁带模拟）。这些反论中最认真的一个是，为了实际使我们相信P_k （k）的真理性，我们应该必须知道该数学家的算法到底是什么，而且必须说服我们，它对取得数学真理的方法有效。如果该数学家在脑中使用一种非常复杂的算法，那么我们就没有机会实际知道这种算法，也就不能实际建立哥德尔命题，更不用说相信它的有效性了。这类反论经常被提出来对抗像我现在要提出的主张，即哥德尔定理指出的，人类的数学判断是非算法的。但是，我自己认为这种反对不能令人信服。此刻我们暂且假定，人类数学家形成其意识判断数学真理的方法的确是算法的。我们将使用哥德尔定理推导出一种荒谬的结果（反证法！）

我们必须首先考虑下列可能性，即不同的数学家使用不等效的算法来决定真理。然而，数学命题的真理性实际上可用抽象的论证决定，这是数学（也许是唯一的学科）最令人印象深刻的特征！假定一个数学论证不含错误，当它完全被理解时，若能使一位数学家信服，就同样能使另一位信服。这也适用于哥德尔型的命题。如果第一位数学家准备接受一个特定形式系统中所有的公理和步骤法则只能给出真的命题，那么他[69] 也应该准备接受这系统的哥德尔命题是描述一道真的命题。这对第二位数学家也完全相同。关键在于，建立数学真理的论证是可传递的[2]。

因此，我们不是在谈论盘旋于不同的数学家头脑中各种难解的算法。我们是在谈论一个普适的形式系统，它等效于所有不同数学家用来判断真理的算法。永远不可能知道，这个假想的“普适”系统或算法是不

是数学家用来决定真理的那一种！因为如果能知道，那我们就能建立起它的哥德尔命题，并且知道那也是数学真理。这样，我们被迫得出结论，数学家实际上用以决定数学真理的算法是如此复杂难解，使得我们永远不知道其有效性。

但这违反数学的宗旨！我们数学传统和训练的主旨是不向我们无望理解的法则权威低头。至少在原则上我们必须了解，一个论证的每一步都能分解成简单明白的步骤。数学真理不是可怕复杂而且其正确性超出我们理解能力的教条。它是从如此简单明白的要素建立起来的，而且当我们理解这些要素时，它们的真理性一目了然，并且所有人都会同意。

按照我的想法，在缺乏一个真正的数学证明时，这是我们能期望得到最明白的反证法！其含义应该非常清楚。数学真理不是我们仅仅用算法决定的东西。我相信意识是我们赖以理解数学真理的关键因素。我们必须“看见”数学论证的真理性，它的有效性才能使人信服。这种“看见”正是意识的精髓。每当我们直接知觉到数学真理时，它就应该呈现。当我们使自己相信哥德尔定理有效时，我们不仅“看见”了它，而且在这么做之时，我们揭露了“看见”过程本身非算法的性质。

灵感、洞察和创造性

我应该对偶尔闪现的新洞察（我们称作灵感）作一些评论。这些思想以及想象是神秘地从无意识的精神中来呢，还是在重要意义上是意识本身的产物呢？人们可以引用许多思想家记载的这类经验。作为数学家，我特别关心其他数学家灵感和创见的思想。但是我想象，在数学和其他科学与艺术中有许多共通之处。我介绍读者阅读非常杰出的法国数学家J·阿达马写的一本薄书《数学发明的心理学》，这是一本非常优秀的经典名著。他引用了许多著名数学家和其他人描述灵感的经验。其中最著名者是由亨利·庞加莱提供的。庞加莱首先描述他着意寻求他称为弗希函数一段紧张的努力，结果陷入绝境。然后：

……我离开我从前居住的坎城，继续进行矿业学校主办的地质学术考察发现。这次旅行使我忘怀自己的数学研究。一到达康坦斯，我们要登上去别的什么地方的公共汽车。正在我的脚踏上阶梯的那一瞬间，与先前的思路毫不相关地，我忽然得到一个发现：我用来定义弗希函数的

变换和非欧几何中的变换完全一样。我没有证实这个思想。我坐在汽车里继续原先开始的交谈，那时没有时间去证实，但是我觉得十分确定。在我回坎城的归程中，我利用空闲之便把它证实了。

这个例子（以及其他许多阿达马引用的例子）的惊人之处在于，庞加莱在一闪念之间得到了这个复杂而高深的思想，那时他的意识思维正专注于完全不同的地方，而且在获得这一思想时还肯定感觉它是正确的，正如后来计算所证明的。应该明白指出，这个思想并不容易用言词解释清楚。如果为了使专家明白这思想，我想他需要做大约一个钟头的学术报告。很明显，就是因为庞加莱先前已经有许多钟头蓄意的意识活动，使他完全熟悉手中问题许多不同的角度，这个思想才能完全成形地进入他的意识。然而，在某种意义上来讲，当庞加莱登上车之时所得到的，是在一瞬间内能被完全理解的“单个”思想。庞加莱确信其思想的真理性更令人惊奇，因此他后来仔细的验证几乎是画蛇添足。

也许我应该试用我自己相仿的经验来做比较。事实上，我想不起来我得到过像庞加莱那样完全从天外而来的妙思异想。（或像其他许多被引用的真正灵感的例子。）我自己是必须有意识地思考手中的问题，也许思考得很含糊，但也许脑中正处于低水平的意识。也可能我正进行其他精神相当放松的活动；例如刮胡子即是一个好例子。也许我刚好开始思考搁置了一段时间的问题。认真进行许多小时从容而清醒的活动肯定是必要的，而且有时我要花一段时间才能重新熟悉一个问题。但是，我也有过经验在“瞬间”得到思想，同时强烈感觉到它是正确的。

也许值得提到一个与此相关的特别奇怪的有趣的例子。1964年秋天，我正为黑洞奇性问题感到烦忧。奥本海默和斯尼德在1939年指出，大质量恒星完全球形的坍缩会导致一个处于中心的空间——时间奇点。广义相对论的经典理论在该处失效（见第7章422页、428页）。许多人觉得，如果他们取消完全球对称（不合情理的）的假设，则这种不愉快的结论就可避免。在球形的情况下，所有坍缩的物质都指向一个中心点，或因为这种对称，所以发生了具有无穷大密度的奇点是可以预料得到的。假设没有这样的对称似乎更合理些，物质以更混乱的方式到达中心区域，不会产生无穷大的密度。或许物质甚至会重新旋转出来，产生与奥本海默和斯尼德理想化黑洞完全不同的行为[3]。

由于新近（20世纪60年代初）发现了类星体，人们重新对黑洞问题

感兴趣，也因而激发了我的思想。这些遥远天体的物理性质使有些人猜测，它们的中心可能是类似奥本海默——斯尼德黑洞的东西。另一方面，许多人又认为奥本海默——斯尼德球形对称假设也许提供了完全误导的图像。然而，从处理另一个问题的经验我想到也许会有一道待证明的数学定理（根据标准的广义相对论）证明空间——时间奇点是不可避免的，并因此证明黑洞的图像必须成立，只要坍缩达到类似“无归点”的条件。我不知道“无归点”（不用球形对称）有任何数学定义的判据，更别说陈述或证明一个适当的定理了。一位同事（厄弗·罗宾逊）从美国来访；当我们沿街走向我在伦敦比尔克贝克学院的办公室时，正滔滔不绝谈论一个完全不同的论题。我们的交谈在跨越人行道时停止了一瞬间，到了另一边又重新开始。就在这短暂的时刻，我显然得到一个思想，但是因为恢复交谈而把它在我脑中遮盖了！

当天，在我的同事离开之后，我回到自己的办公室。我记得有种古怪又难以解释的兴奋感觉。我开始把整天在我脑袋里发生的所有事情都想过了一遍，试图找出引起这种感觉的原因。在排除了许多不足够充分的原因后，最后想起了我跨过马路时得到的想法。这想法为我头脑中琢磨许久的问题提供解答，并使我一瞬间欣喜万分。这想法显然正是我需要的判据，后来我将之称作“捕获面”。然后，没花很长时间我就得到了寻求中的定理证明概要（Penrose　1965）。尽管如此，我花了一段时间才把该证明以完全严格的方式写出，但是我穿越街道时所得到的思想是一个关键。（有时候我怀疑，如果那天我还经历了其他不重要的事，我也许就根本记不得捕获面的想法！）

上面轶事使我想到另外有关灵感洞察的论题，就是我们在形成判断时，美学标准具有重大价值。美学标准对艺术来说是至高无上的。在艺术中美学是门高深的课题，哲学家们奉献终身去研究它。可以说在数学和科学中，美学标准仅是偶然的，而真理标准才是至高无上的。但是在人们考虑灵感和洞察问题时，似乎不可能把两种标准分开。我的印象是，坚信瞬间灵感是正确的（我应该加一句，并非百分之百可靠，但至少比纯粹碰运气可靠得多）与灵感的美学品质有很密切的关系。看起来漂亮的思想比看起来丑陋的思想对的机会更大得多。这至少是我自己的经验，其他人也表达过类似的感想（参阅*Chandrasekhar　1987*　）。例如，*Hadamard 1945*，第31页）中写道：

……很显然若没有探索的意志，任何有意义的发现或发明都不会发

生。但是我们在庞加莱的经验中看到了一些别的什么，美感的干涉作为一个不可或缺的探索手段。我们得到了两重结论：

发明是一种选择。

这种选择绝对是由科学的美感所控制的。

例如还有狄拉克（1982）毫不羞愧地声称，正是他敏锐的美感使他预知电子的方程式（指的是369页的“狄拉克方程”），而其他人却无法找到。

我自己的思维肯定可以证明美学品质之重要性，不管是指一种可以“坚信”的“灵感”，或是当一个人朝期望目标摸索时必须持续进行的一种“例行”猜测。我曾在别处写过这相关的论题，特别是有关图10.3和4.11描述的非周期性镶嵌。毫无疑问，这些镶嵌中第一个镶嵌的美学品质——不仅是它的视觉外观，还有它迷人的数学性质——给了我一种直觉（可能是在“一瞬间”，但是大约只有60%肯定！），它可以由合适的搭配规则（也就是锯齿式组合）排列出来。我们很快就要再看到这些镶嵌模式（参阅*Penrose 1974* ）。

我对这一点是非常清楚的，美学标准的重要性不仅适用于灵感的瞬息判断，而且也适用于我们在数学（或科学）研究中更须经常作的判断。严格的论证通常是最后的步骤！人们在此之前必须作许多猜测，美学信仰对于这些是极重要的，它总是受逻辑论证和已知事实的约束。

我正是把这些判断当成意识思维的标志。我猜想，即使是突然闪现的灵感，很显然也是由无意识的精神现成准备好了的。意识正是裁决者，如果思想不是“听起来不错”的话就会很快地被否决并忘掉。（古怪的是，我实际上的确忘记了我的捕获面，但是这不在我所指的同样水平的忘记。该思想进入意识的时间足够长，因而留下永久的印象。）我是在假定，我所指的“美学”否决是完全禁止没有魅力的思想到达意识的相当永久的层次。

那么在我的观点中，无意识在灵感思维中的作用是什么呢？我承认，这题目不像我希望的那么清楚。无意识似乎的确在这范围扮演重要的角色，我应该同意一个观点，无意识过程很重要。我还应该同意，无意识的精神绝非仅仅随机地吐出思想来。必须存在一种强有力的选择步

骤，使得意识精神只受“有机会的”思想扰动。我提议，这些选择判据（多半是“美学”的）已经被意识迫切的希求所影响（正如数学思想和已经建立的原理不协调时，就会有丑恶感觉伴随而来）。

与此相关的问题是，什么才构成真正的创造性。我觉得它牵涉到两个因素，也就是“提出”和“淘汰”过程。我想，“提出”过程大多是无意识的，而“淘汰”过程大多是有意识的。缺少有效的提出过程，根本就不会有新思想。但若仅有提出过程，则它的价值非常小。人们需要一个形成判断有效过程，使只有具备合理成功机会的思想留存下来。例如在睡觉中非常容易涌现奇思异想，但是很少能在清醒意识的严厉批判下存活下来。（我本人在睡梦状态就从未得到过成功的科学思想，而别人就幸运得多，譬如化学家凯库勒，发现了苯结构。）依我的意见，是意识的淘汰过程（也就是判断）而不是无意识的提出过程作为创造力的问题中心，但是我知道许多人持相反的观点。

在离开这令人相当不满意的状态之前，我应该提到灵感思维的另一个特点，就是它的全局特征。上述的庞加莱轶事是个显著的例子，在极短暂的时间内来到他头脑中的思想包含了大量的数学思维。非数学读者也许更能立即接受，艺术家（某些）把他们的创作整体一下子放到头脑中（虽然毫无疑问一样地难理解）。莫扎特（正如*Hademard 1945* ，第16页所引用的）生动地提供了一个使人惊奇的例子：

当我感觉良好或处于风趣状态时，或者当我在美餐后驾车兜风或散步时，或者在难以入眠的夜晚，思绪犹如潮水般地涌进我的头脑。它们从何而来又如何来呢？我不知道，这与我无关。我把那些喜欢的留在脑中，并且轻轻地哼唱；至少别人曾告诉我，我是这么做的。一旦我得到了主旋律，其他曲调就依照整个乐曲的需要连接进来和旋律配合，最后每一种乐器的配乐以及所有的旋律片断也参与进来，最后就产生了一部完整的作品。此时灵感在我的灵魂中燃烧。作品渐渐成熟，我不断地扩展它，把它孕育得越来越清晰，直到整个曲子在我的头脑中完成，尽管它可能很长。该乐曲在我的精神中正如一幅美丽的图画或一位英俊的少年在眼前闪现。它不是连续地来到我的头脑中，而是我的想象使我完整地听到它，然后才完成细节部分。

我觉得这和提出/淘汰的方案一致。虽然“提出”无疑是极有选择性的，不过它看来是无意识的（“它与我无关”）；而淘汰是有意识的品味

仲裁人（“我把那些喜欢的留在脑中……”）。灵感思维的整体性在莫扎特的引语中特别明显（“它不是连续来到……而是完整地”）。也正如庞加莱的例子一样（“我没有证实这个思想。……那时没有时间去证实”）。此外，我还坚持，一般来说，我们意识思维已经呈现出明显的整体性。我将很快回到这个问题上来。

思维的非言语性

阿达马研究创造性思维令人印象深刻的要点之一，是拒绝接受迄今仍常听到的论题——言语是思维所必需的。引用爱因斯坦致阿达马信中的一段话就能把这问题解释得再好不过了：

词语或语言，无论是写的或说的，在我的思维机制中，似乎都不起任何作用。似乎作为思维要素的精神实体是一些能够“自动”复制与结合的特定符号和一些大致还算清晰的图像……在我的情形中，上面提到的要素有视觉的和肌肉之类型。只有在第二阶段，当所提到的联想活动充分建立起来并能随意复制时，才须费心寻找习惯的词语或其他符号。

杰出的遗传学家佛朗西斯·盖勒顿的一段话也值得引用：

写作是我的严重缺陷，言语表达的缺陷更严重。我用语言方式来思考比用其他方式更不容易。经常发生这样的事，在经过辛苦的工作后得到完全清楚和满意的结果，但当我试图用语言来表达时，我必须先使自己位于另一个完全不同的智力层面。我必须把自己的想法翻译成和它们不甚配合的语言。因此我在寻求合适的词汇和短语中浪费了大量的时间。我意识到，当突然必须演讲时，经常仅因为言语笨拙而不是因为缺乏清楚认知，使得我的演讲变得非常难懂。这是我生活中的小烦恼。

阿达马自己也写道：

我坚持，当我真正进行思考时，词语在我的头脑中根本不存在。我的情形和盖勒顿完全一样。甚至在读到或听到一个问题后，从我开始思考的那一时刻起每一个词都消失了；我完全同意叔本华所写的：“思想一旦被语言具体化就马上死去”。

因为这些例子和我自身的思维模式非常一致，所以才在这里引用。

我几乎所有的数学思维都是按照视觉以及非语言的概念进行的，虽然这种思维经常伴随着愚笨并且几乎无用的言语评论，诸如“这件事跟着那件事，而那件事又跟着另一件事”。（我有时在简单的逻辑推导中会用到语言。）还有，我还经常亲身体验到这些思想家把他们思想翻译成语言时所遭遇到的困难。经常的原因就是找不到言语来表达需要的概念。事实上，我时常利用特别设计的图表来计算（参阅*Penrose and Rindler 1984*，424—434页），这些图表代表某类代数表达式的速记。把这些图表翻译成文字会是非常烦琐的过程，只有在必要向他人仔细解释时才把它当成最后的手段。还有一个相关的观察是，我曾注意到，如在我潜心于数学时，有人忽然要和我交谈，我在几秒钟内几乎不能说话。

我不是说我从来不用语言方式思考，只是我发现语言对数学思维几乎没有用处。其他种类的思维，譬如哲学也许更适合于用言语表达。这大概是为什么许多哲学家抱持一个观点认为，语言是智力或意识思维的根本！毫无疑问，不同的人以不同的方式思考，甚至仅就不同的数学家而言——这是我自己的经验。数学思维的主要倾向可分为解析式和几何式。虽然阿达马用视觉图像而不用言语描述来进行数学思考，但有趣的是他本人认为自己是用解析的方式思考。至于我自己则是非常倾向用几何方式思考。但是，各个不同的数学家的思考倾向的范围非常广阔。

一旦接受大多有意识思维确实具有非言语的特征——依我看来，基于前面一些理由这个结论是不可避免的——那么也许读者不难相信意识思维也具有非算法的成分！

记得在第9章（483页）我提到一个屡见不鲜的观点，只有具有语言能力的那一半头脑（绝大多数人是左半边），可有意识能力。按照上面的讨论，读者应该很清楚为何我发现这种观点完全不能接受，我不知道总体来说，数学家是否倾向利用头脑之一半比另一半更多；但是毫无疑问，真正的数学思维须有高水平的意识。解析思维主要是在左半脑进行的，而几何思维通常归于右半部，所以可以很合理猜测大量有意识的数学活动实际上发生在右半边！

动物意识

在结束言语化对意识的重要性这个论题之前，我要讨论早先曾简要

提起的问题：非人类动物能否有意识。我觉得人们有时依据动物不能言语来推断它们不具备任何可觉察的意识，而且隐含着反对它们具有任何“权利”。读者容易看出，我认为这种论证是站不住脚的。这是因为很多复杂的（例如数学）意识思维不用言语就能进行。还有右半脑有时被认为只有和黑猩猩一样“少”的意识，亦是因为黑猩猩缺少言语能力（参阅*LeDoux 1985*，197页—216页）。

事实上，当允许黑猩猩或大猩猩使用符号语言，而不用正常人类的方式讲话时（它们不能讲话是由于缺少适用的声带），它们是否真正有言语能力引起许多争议（参阅*Blakemore and Greenfeld 1987* 的各种文章）。不过争议归争议，清楚的是，它们使用这种方法至少在某些基本程度上能互相沟通。依我自己的意见，有些人不承认这方式为“言语”是有点过于吝啬，也许有些人希望借口拒绝让猩猩进入言语俱乐部，因而排除它们进入有意识生物的俱乐部！

先不管语言的问题。有很好的证据显示黑猩猩能有真正的灵感。昆拉德·洛伦斯（1972）描述过一只关在房间里的黑猩猩，一根香蕉悬挂在天花板，刚好使猩猩拿不到，并且在房间其他地方放一个盒子：

这事使得它烦躁不安，它又回到那里去。然后——没有更佳方式可以描述——它原先阴郁的脸忽然“发亮起来”。现在它的眼光从香蕉移到香蕉正下方的空地，从这里移到盒子那里，又移回空地来，再移到香蕉那里去。下一刻，它欢呼了起来，以极其高昂的情绪翻一个筋斗到了盒子旁边去。它把盒子推到香蕉下面，完全确信自己会成功。所有看到这一幕的人都不会怀疑类人猿体验到真正的“灵光一现”。

注意，正如当庞加莱踏上公共汽车时所经验的那样，黑猩猩在证实它的思想之前就“完全确信成功在握”。我认为这种判断需要意识。如果我是对的，那么这里就有证据显示非人类动物的确有意识。

有关海豚（及鲸鱼）还产生了一个有趣问题。人们会注意到，海豚的大脑和我们的一样大（甚至更大），海豚还能相互传递极其复杂的声音信号。也许它们为了于人类尺度或近似人类尺度的某种有别于“智慧”的目的而需要相当大的大脑。而且，由于它们缺乏适于抓、拿的手，不能建造我们能鉴赏的这种“文明”。虽然为着同一原因，它们不能写书，但或许它们有时像哲学家，沉思生活的意义以及为何它们在“那

里”！它们是否有时通过复杂的水底声音信号来传递它们的“知觉”呢？我不晓得有任何研究指出它们是否用头脑特定的一边来“言语”并相互沟通。在和施行于人类的“分裂头脑”手术以及所隐含的“自我”连续性这令人困惑的意义相关联的方面，我们应该提到海豚不是整个头脑同时进入睡眠状态 [4] ，而是每次只有一半头脑睡着。如果我们能询问它们对意识的连续性有何“感觉”，那将会很有教益！

与柏拉图世界的接触

我提到过，不同的人似有许多不同的思考方式，而且不同的数学家也以不同的方式思考数学。我记得当我将要进大学研习数学时，以为会发现我未来的数学界同行多少会用和我一样的方式思考。以我在学校的经验是，我的同学思考方式似乎和我很不同，这使得我有点受挫。我本来兴奋地以为：“这下我可以找到很容易交流的同道了！有些人的思考方式比我的更有效，有些人差一些，但是所有人的脑波频率都和我一样。”我大错特错了！我相信，我比以前经验到更多的不同的思考模式！我的思考方式比他人较多几何成分而较少解析成分，但是我其他同事的思考模式有许多其他差异。我对于理解一个用言语方式解说的公式总是感到困难，而我许多同事似乎毫无这种困难。

当一位同事想对我解释一段数学时，通常我的经验是，我必须全神贯注地听，但是对一组词和另一组词之间的逻辑关联几乎完全不能理解。然而，在我脑中会形成一种猜测图像代表他所要传达的思想。这个图像完全是按照我自己的方式形成，而且和我同事所理解的脑中图像关系不大。经过这过程之后，我才能回答。令我相当吃惊的是，我的评语通常被接受，而交谈就以这种方式来回进行下去。在交谈结束时可以很清楚地看出，确实进行了一种真正而正面的交流。然而我们各自呢喃的实际句子似乎只有少数时候能被真正理解！在我成为专业数学家（或数学物理学家）之后这些年，我觉得这种现象比我当大学生时更为显著。也许随着我的数学经验增加，使我更容易猜测他人的解释想表明的意义，也可能使我自己解释事物时更能容忍其他的思考模式。但是在本质上并没有什么改变。

我自己经常感到困惑，按照这种奇怪的步骤如何能沟通。现在我想大胆提出一种解释，因为我认为它可能和我曾讨论过的其他问题有很深

的关联。关键在于，人们在讨论数学时不只是传递事实。从一个人向另一个人传达一连串（偶然的）事实时，第一个人必须把所有事实仔细说明，而第二个人必须一一吸收进去。但是对于数学而言，事实的内容非常少。数学的陈述必须是真理（否则便是谬误！），即使第一位数学家的陈述仅是探索这样一个必要的真理，假定第二位充分理解前者的陈述，那么正是真理本身被传达给第二位数学家。第二位的脑中图像也许在细节上和第一位的图像不同，他们的言语描述也可以不同，但是相关的数学思想则在他们之间共通交流。

若不是有趣或高深的数学真理在一般数学真理中寥若晨星的话，则这类沟通根本不可能。譬如，要沟通像4 897×512=2 507 264这样乏味的陈述，在第一位表达这精确的陈述时，第二位的确必须要能理解前者。但是，对于数学中有趣的陈述，即使描述非常不精确，人们经常仍然能够掌握所要传递的概念。

由于数学是精确度最高的学科，这里似乎存在一个佯谬。的确，在书面上为了保证各种陈述既精密又完整，人们必须十分费心。然而，为了传达数学思想（通常利用言语描述），这种精确性有时会先产生抑制作用，而可能需要更模糊的叙述性传递形式。在掌握了观念的实质后再考虑细节。

数学观念如何能用这种方式传递呢？我想只要头脑在感知一个数学观念，它就是和数学观念的柏拉图世界接触。（回想一下，按照柏拉图的观点，数学观念本身是存在的，它存在于柏拉图的理想世界里，只有通过智慧才能接触到，参阅127页，205页。）当有人“看见”了一个数学真理，他的意识突破到这个理念世界中去，并与之直接接触（“通过智慧来接触”）。我描述过这种“看见”与哥德尔定理的关系，而它是数学理解的精髓。正是由于每位数学家都有直接通往真理的道路，他们之间的相互交流才有可能。每一个生物的意识都是通过这个“看见”的过程，来直接感知数学真理。（的确，这种感知的行为时常伴随着“啊，我看到了！”的惊喜！）由于每人都能和柏拉图世界直接接触，他们比人们所预期的更容易进行交流。当进行这种柏拉图接触时，各人在每种情形下所具有的精神图像也许相当不同，但是由于大家直接和同一外部存在的柏拉图世界接触，所以才可能进行交流！

按照这种观点，精神总是能够进行这种直接接触。但是每一次只能

进行一点。数学的发现包含接触范围的扩展。由于数学真理必须是真理，在技术的意义上讲，并没有实际的“信息”传递给发现者。所有信息一直存在那里。人们只不过是把东西放在一起并“看见”了答案！这和柏拉图自己的观念非常一致，发现（譬如数学）只不过是一种记忆形式！的确，我就经常感到吃惊，因为记不住某人名字和找不到正确的数学概念之间具有相同点。在每一种情形下，所要寻找的概念在某种意义上已经存在我的脑中，尽管尚未发现的数学观念具有更不平常的语言形式。

为了使这种观察事物的方式有助于数学交流，人们必须想象，有趣高深的数学观念比乏味平凡的思想更可能存在。这对于下一段猜测性考察具有重大意义。

物理实在的一个观点

意识如何能在物理实在的宇宙中产生，任何有关的观点至少要包含解决物理实在本身的问题。

例如，强人工智能的观点认为，“精神”通过一个足够复杂的算法体现找到了自身的存在，而这个算法可由物理世界的某物体来执行。而这些实际的物体究竟是什么并没关系。神经信号、沿着导线的电流、齿轮、滑轮或水管都可做得一样好。算法本身被认为是所有关键之处。但是，对独立于任何特殊的物理体现而“存在”的一个算法，柏拉图的数学观点似乎是必要的。一位强人工智能支持者很难采取不同观点，如“数学观念只存在于精神中”。因为这会导致逻辑循环，为了算法的存在，预先需要精神的存在，而为了精神存在，则预先需要存在的算法！他们也许企图采取这样的论证，即算法可作为一张纸上的痕迹、一块铁上的磁化方向或一台电脑存储器上的电荷位移而存在。但是，这种物质形态自身实际上不具有算法。为了得到算法，它们需要一个解释，也就是必须能对这些形态解码；这就要依赖写这算法的“语言”。为了理解这语言，预先存在的头脑似乎又是必需的，这样我们又回到了出发之处。那么，我们就接受算法处于柏拉图世界中。根据强人工智能的观点，那个世界正是精神之所在。我们现在就必须面对物理世界和柏拉图世界如何相互关联的问题。依我看来，这正是强人工智能对精神——身体问题的说法！

既然我相信精神（意识的）不是算法实体，我自己的观点与上述不同。但是，当我发现在强人工智能和我自己的观点之间有许多共同点时有些受窘。我曾指出，我相信意识和必要真理的感知有密切关联，并因此得以和柏拉图的数学概念世界直接接触。这不是一个算法的过程，我们并不特别关心也许栖息在那个世界的算法。但是根据这个观点，再一次看到精神——身体问题密切关系着另一个问题：柏拉图世界与具有实在物体的“真实”世界如何相关。

我们在第5章和第6章看到，实际物理世界以惊人方式符合一些非常精密的数学方案（参阅197页的超等理论）。人们经常评论这些精密度是何等不寻常（尤其参阅*Wigner 1960*）。我很难相信光靠随机自然选择加以淘汰，使得只有好的思想保存下来，就能产生超等的理论，像有些人企图坚持的。好的思想实在是太好了，用这种随机淘汰后留存的方式根本不可能产生。必须有一种更深入的基本原因使数学和物理之间，也就是柏拉图世界和物理世界之间相符合。

就“柏拉图世界”而言，人们赋予了它某种实在性，可以在某方面和物理世界的实在性相比。另一方面，物理世界本身的实在性显得比发现相对论和量子力学的超等理论之前更加模糊了（参阅197页、198页，尤其是367页的评论）。正是这些理论的精确性为实际物理实在提供近乎抽象的数学存在，这难道不是一个佯谬吗？具体的实在怎么会变成抽象和数学的呢？这也许是抽象数学概念如何在柏拉图世界中获得近乎具体实在的反面问题。也许就某种意义来说，这两个世界是同一的？（参阅*Wigner 1960*；*Penrose 1979a*；*Barrow 1988*；还有*Atkins 1987*。）

虽然我强烈同情实际上把两个世界视为同等的这种思想，对这问题还有更多讨论余地。正如我在第3章和本章前面提到过，某些数学真理比其他的具有更强烈的（“更深刻的”、“更有趣的”、“更富有成果的”？）的柏拉图实在性。这些也就更强烈等同于物理实在的运行。[复数系统（参阅第3章）就是一个例子，它是量子力学的基本部分，即概率幅。]利用这种认同性，“精神”如何能揭示出物理世界和柏拉图数学世界之间某种神秘的联结就更容易理解。我们还可回忆在第4章描述过，数学世界中有许多部分，而且有些是最深奥最有趣的部分，有非算法的特性。所以，在我试图详细解释的观点基础上，非算法行为很可能在物理世界中具有非常重要的作用。我设想，这种作用和“精神”的概念本身密不可分。

宿命论和强宿命论

迄今为止我对于“自由意志”的问题讲得很少，自由意志通常被当作精神——身体问题主动部分的基本论题。我的精力集中于设想意识行为的作用本质上有非算法的一面。我们记得，在大多数超等理论中存在一种清清楚楚的宿命论。就这种意义来说，如果我们知道系统在任一时刻的态 [5] ，那么理论的方程式把该系统的态在以后（或以前）的任何时刻完全地固定死。由于一个系统未来的行为似乎被物理定律所完全决定，因此似乎没有任何“自由意志”的余地。甚至量子力学的U部分也具有这种完全决定性的特征。然而“量子跃迁”R不是宿命论的，它把完全随机的因素引进时间演化中来。早先，许多人踊跃接受以下可能性，即这里可以是自由意志用武之地，意识的作用对单独系统跃迁的方式也许有某种直接效应。但是，尽管我们希望我们的自由意志有所作为，如果R是真正随机的，则它也不会有多大帮助。

虽然我的观点在这方面尚未很明确，不过我认为有些新过程（CQG；参阅第8章）可能超越在U和R（现在这两者都被认为是它的近似）之间的量子——经典界限，而这个新的过程包含本质上非算法的因素。其中一个含义是，甚至即使未来可以被现在所决定，它也不能从现在计算出来。我在第5章的讨论中试图清楚地把可计算性从决定性中区别出来。我以为CQG是决定性但非计算性的理论很可以说得通 [70] 。（回忆一下我在第5章220页所描述不可计算的“玩具模型”。）

人们有时采取这样的观点，即使是经典（或U—量子）宿命论也不是一个有效的宿命论，因为不能真正充分知道初始状态，使得将来实际上能被计算出来。有时初始条件非常微小的改变会导致最后结果非常大的差异。例如发生在（经典的）宿命性系统中被称作“混沌”的现象——天气预报的不确定性即为其中一例。然而，非常使人难以置信的是，这种经典的不确定性会允许我们的自由意志（或只是幻象？）。虽然我们不能计算出未来的行为，但是一直从大爆炸开始，未来行为仍然是被决定了的（参阅225页）。

同一个反对意见也用来反对我的建议。从这个观点看，未来世界虽然不是可计算的，但仍然被过去所完全固定，这可以一直回溯到大爆炸。实际上，我并非独断地坚持CQG必须是决定性而非计算性的。我猜想我们寻求的理论会比这些描述更加微妙。我只要求这理论必须本质上

包含非算法的因素。

为了结束这一节，我想评论一下人们对宿命论可能坚持采取更极端的观点。这就是我所谓的强宿命论（*Penrose 1987b*）。根据强宿命论，不仅未来的事由过去所决定；根据某种精密的数学方案，宇宙在所有时刻的全部历史都是固定的。因为柏拉图世界是一下子就全部固定好了的，对这宇宙并没有什么“其他可能性”！如果人们倾向于认定柏拉图世界和物理世界相同，这种观点颇具魅力。（我有时怀疑，当爱因斯坦写下“我所真正感兴趣的是，上帝是否能以不同的方式来创造世界；也就是说，必要的逻辑简单性是否为自由选择留下任何余地！”时，不知在他脑中是否有过这种方案。（致恩斯特·斯特劳斯的信；见*Kuznestsov 1977*，P285）。

人们可以把量子力学的多世界观点（参阅第6章378页）当作一个变种的强宿命论。根据这类观点，一个精确的数学方案固定的不是单独的个别宇宙历史，而是固定了所有无数个由它所决定的“可能的”宇宙历史。尽管这个方案（至少对我来说）呈现出令人不满意的性质和一大堆问题与缺陷，我们却不能排除这方案的可能性。

我觉得，如果人们持强宿命论但同时不持多世界观点，则制约宇宙结构的数学方案就可能必须是非算法的[6]。原因在于，如果不是这样，人们便可以原则上计算出下一时刻将要发生的事，然后他可以“决定”去做其他完全不同的事，这就会在“自由意志”和这理论的强宿命论性质之间产生显著矛盾。在理论中引进不可计算性就会避开这一矛盾——虽然我必须承认，我对这种解决办法颇感不舒服，而且我还预料，有些更加微妙的、实在的（非算法的！）规则实际在制约这个世界的运行！

人存原理

意识对于整个宇宙有多重要呢？缺少任何有意识的居住者的宇宙能否存在呢？物理定律是不是为了允许意识生命存在而特别设计出来的呢？我们在宇宙的空间或时间中的位置是否有任何特殊的地方？这些就是所谓人存原理所讨论的问题。

该原理有许多形式。（见*Barrow and Tipler 1986* 。）这些讨论中最被广泛接受的仅仅是意识（或“智慧”）生命在宇宙空间——时间中的定位。这是弱人存原理。这种论证可以用来解释，现在地球上的条件为何刚好适合于地球上（智慧）生命的存在。如果条件不是刚好，我们不应发现自己现在处在这个地方，而是在别的什么地方，在其他适当的时间。布兰登·卡特和罗伯特·狄克非常有效地利用此原则解决了困惑物理学家许多年的问题。这问题是关于从观察发现的物理常数（引力常数、质子质量、宇宙年龄等）之间保持的各种令人惊讶的数值关系。令人不解的是，有些关系只有在现代地球历史才成立，所以我们刚好生活在这非常特殊的时期（大概几百万年！）。后来卡特和狄克用下列事实来解释：这个时期同被称为主序星（如太阳）的生命周期一致。在其他任何时期，按照同样的论证，四周就不会有智慧生命来测量讨论中的物理常数，所以这种巧合必须成立，因为只有在这巧合成立的特定时刻四周才会有智慧生命！

强人存原理牵涉得更广。在这情形下，我们不仅关心自己在这宇宙中，也关心在无限个可能的宇宙中空间——时间的定位。我们现在可以回答为什么物理常数或一般物理定律要特别设计才能使智慧生命得以存在。其论证是，如果这些常数或定律是不同的，则我们就不应该处于这个特定宇宙中，而应该处于其他宇宙中！依照我的意见，强人存原理有个可疑的特征，好像只要理论家提不出更好的理论去解释观察的事实，就会提出强人存原理（也就是在粒子物理理论中，粒子的质量是没有解释的，人们因而断言，如果它们的数值和被观察到的数值不同，则生命便不可能存在，等等）。另一方面，假定人们小心地使用弱人存原理，我觉得它是无懈可击的。

由于使用人存原理——不管是强的还是弱的——人们可以尝试展示，由于知觉生物，也就是“我们”，必须存在以观察世界的这一事实，意识的存在便是不可避免的，所以人们不必像我以前一样，去假定知觉具有任何选择优势！我的看法是，这个论证技术上是正确的，弱人存原理的论证（至少）能为意识不需自然选择的帮忙而存在提供原因。另一方面，我相信人存论证不是意识演化的真正（或仅有）原因。从其他方向有足够的证据使我信服，意识具备强而有力的选择优势，而且我认为人存论证是不必要的。

<u>镶嵌和准晶体</u>

我现在要从前几节的大胆猜测转而考虑更科学和更“具体”的问题，虽然仍有一点猜测性。这个问题初看起来有点离题。但是它对我们的意义在下一节就会变得明显。

我们回忆在180页图4.12中的镶嵌模式。这些模式令人惊异之处在于，它们“几乎”违反了一个与晶格有关的标准数学定理。该定理叙述道：在结晶模式中只允许二重、三重、四重和六重的旋转对称。所谓结晶模式，我是指具有平移对称点的分立系统。所谓平移对称是说，用一种自身滑动而不转动的方式，使得该模式和自身相重合（就是说移动不会改变该模式）而且因此有周期性的平行四边形（见图4.8）。图10.2绘出了这些允许的旋转对称的镶嵌模式例子。现在图4.12的模式，正如图10.3中那样（它基本上是由179页图4.11的花砖拼在一起产生的镶嵌），却又几乎具有平移对称和几乎具有五重对称。这里“几乎”的意思是：人们可以找到模式（分别为平移和旋转）的运动，并且这模式的自身重合能达到任何预先指定的比百分之百略低的相合性。我们在此没有必要去忧虑它准确的意思。我们在这里所关心的是，如果有一种物质的各原子被安置在这种模式的各顶点，则这物质就显得像晶体，但它会呈现出被禁止的五重对称性！

1984年12月，正在美国华盛顿首都国立标准局与同事共同研究的以色列物理学家邓尼·谢茨曼宣布发现了一种铝锰合金的相，它的确像类晶体物质，现在称为准晶体，它具有五重对称。事实上，这种准晶体还具有存在于三维中而不仅是平面上的对称性，这就给出了总共有正二十面体的对称性（*Shechtmanetal.1984*　）。（类似我的五重平面镶嵌，三维“二十面体”类似物被罗伯特·阿曼在1975年发现，见*Gardner 1989* 。）谢茨曼的合金只能形成大约千分之一毫米非常微小的准晶体。但是，后来还发现其他的准晶态物质，尤其是一种铝——锂—铜合金，其二十面体对称单元可长成大约1毫米的尺度，用肉眼都能完全看得到（图10.4）。

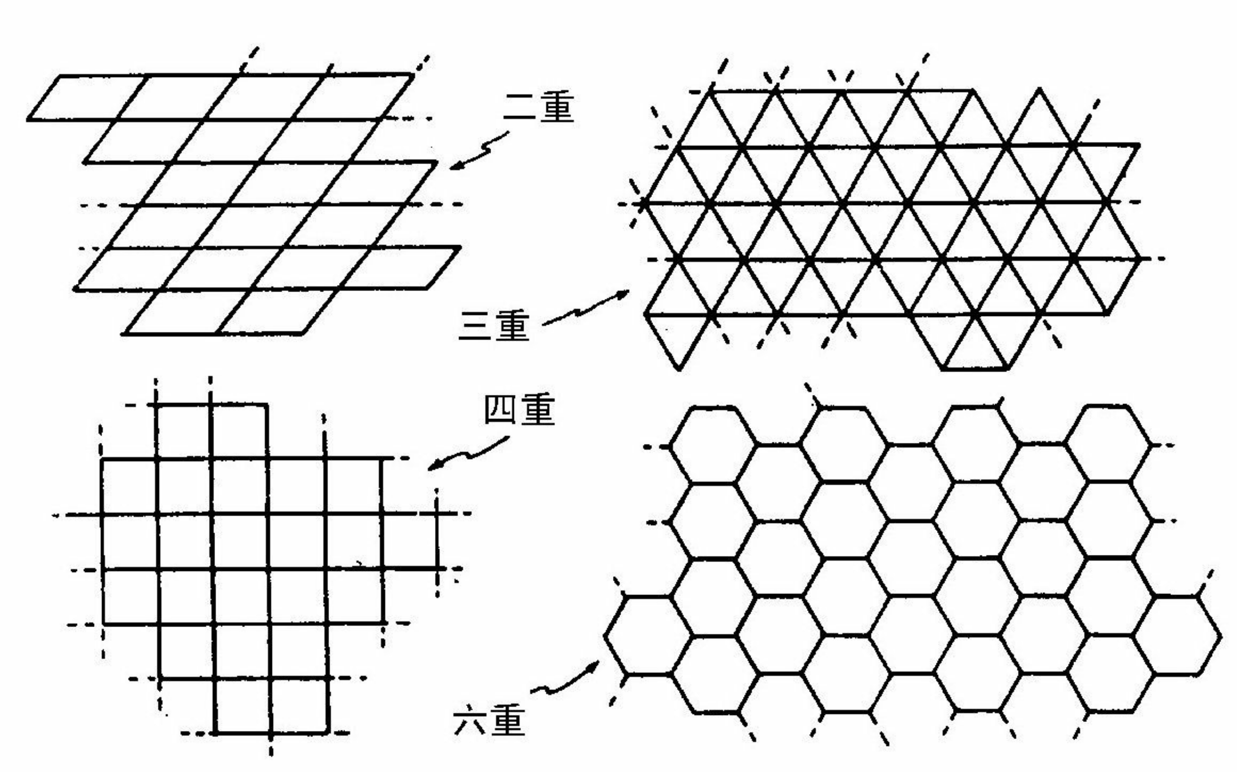


图10.2　具有不同对称的周期性镶嵌（在这里的每一种情形中，把花砖的中心当作对称的中心。）

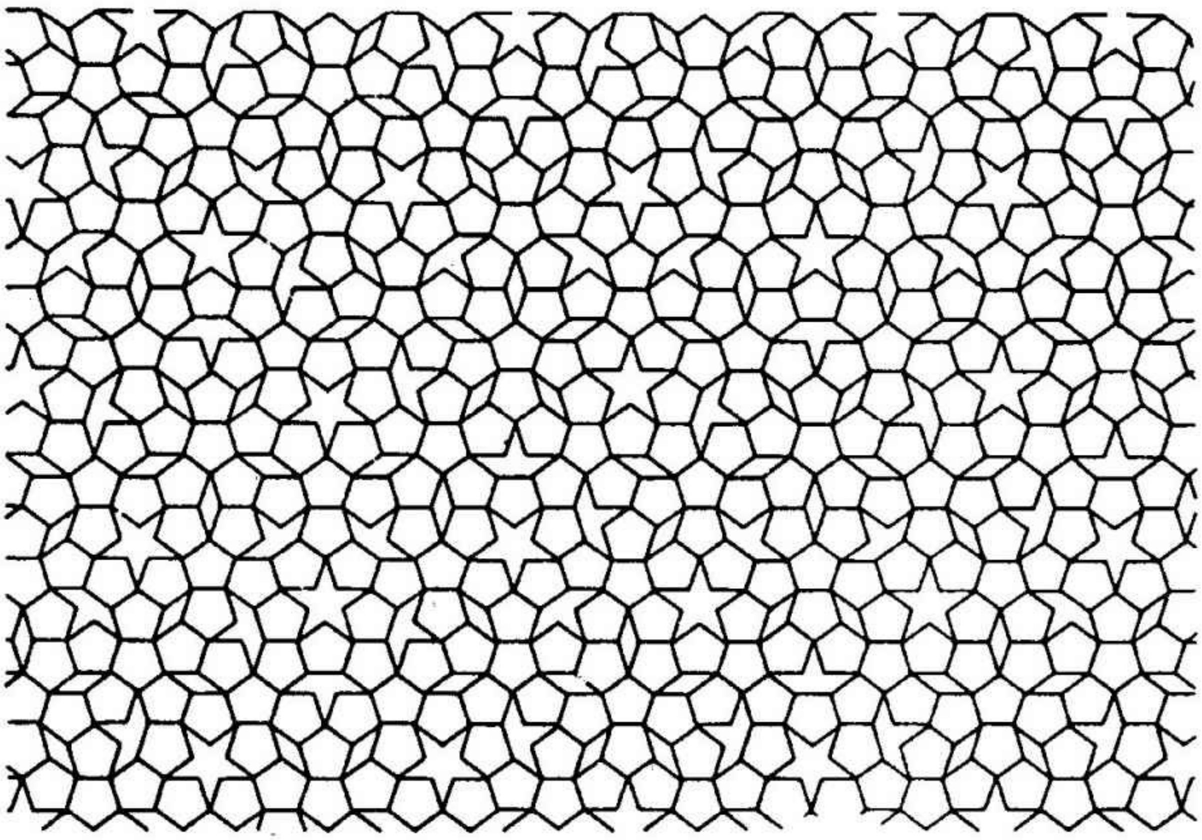


图10.3　一种准周期性镶嵌（基本上是由图4.11中的花砖拼在一起产生的）。它具有晶体学中“不可能的”五重准对称

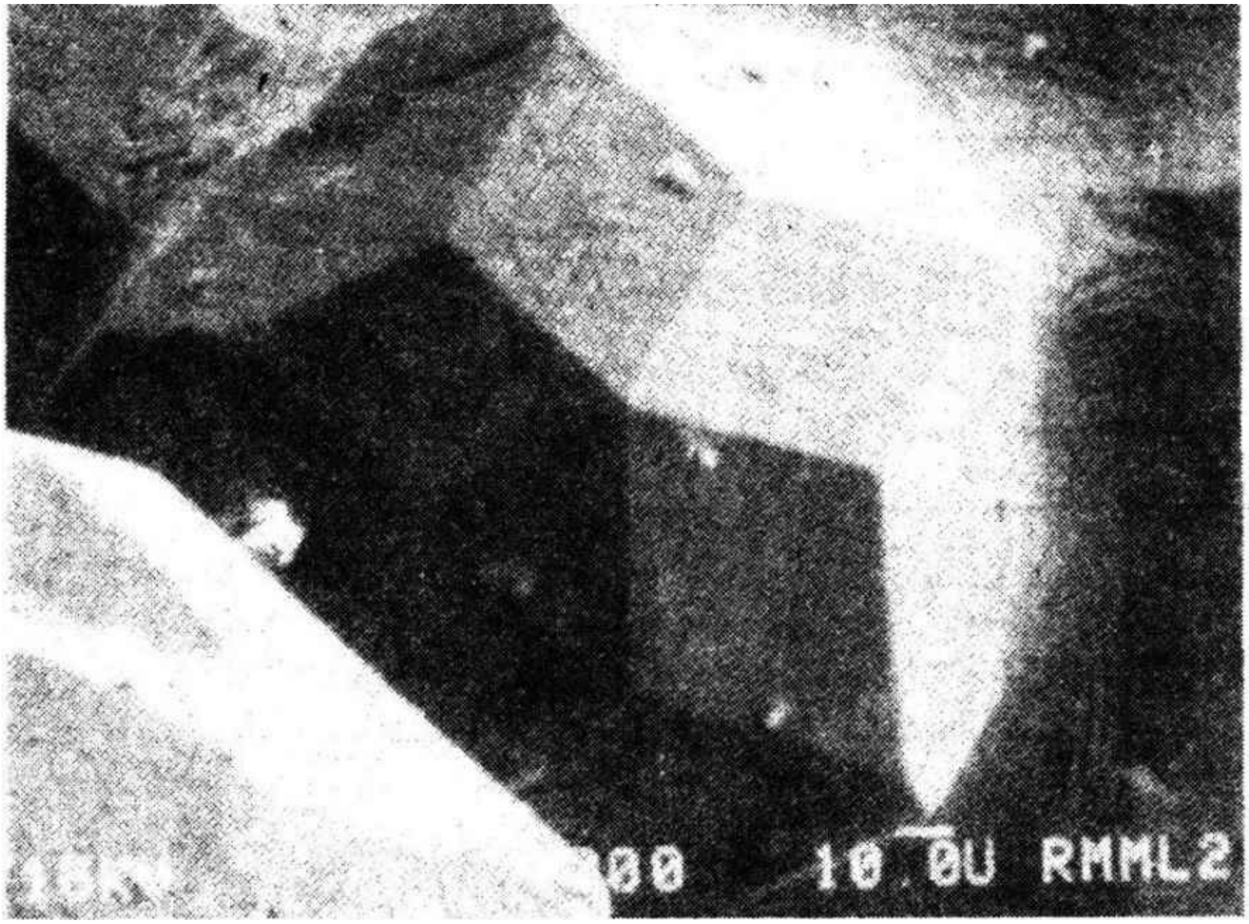


图10.4　这准晶体显然具有不可能的晶体对称性（一种铝——锂—铜合金）（取自盖里1987）

现在我所描述的准晶体镶嵌模式有一个显著特征，即它们的拼合必须是非局部的。这就是说，在装配该模式时，必须不时考察距离装配点许许多多“原子”之遥的模式状态，以保证把许多小块放在一起时不发生严重的错误。（这也许有点像我在讨论自然选择时，曾提及的“智慧探求”。）这一类特征是当前围绕准晶体结构和成长问题重大争议的一部分。在解决一些突出的问题之前，匆忙作确定的结论显然是不明智的。尽管如此，人们可以猜测，而且我将大胆提出自己的意见。首先，我相信这些准晶态物质的确具有高度组织，而且它们的原子排列和我考察过的镶嵌模式相当接近。其次（我的意见是属于比较尝试性的）这意味着不能按照符合晶体成长的经典图像，依靠每次局部地添加一个原子来合理地完成它们的装配，它们的装配一定有本质上量子力学的非局部性因素[7]。

我描绘这种成长的发生方式是，原子不是单独来到并自己附到连续移动的成长线上去（经典晶体成长）。人们必须考虑附加原子的许多不

同排列的量子线性叠加的演化（量子过程U）。这的确是量子力学告诉我们必须（几乎总是）发生的！不是只发生一个原子排列；许多不同的原子排列必须在复线性叠加中共存。这些不同的叠加选择中有一些会成长为大很多的团块，而且在某一点，某些不同选择的引力场之间的差别将会达到单引力子的水平（或不管什么适当的水平；见第8章466页）。在这阶段，其中一种排列（或可能仍是一种叠加，不过却是一种缩减了的叠加）会成为“实在”的排列而被挑中（量子过程R）。这个叠加形态加上更确定的形态减缩会以越来越大的尺度继续下去，直到形成相当尺度的准晶体。

正常情况下，当大自然寻求一种晶体排列时，它总是寻求具有最低能量的排列（把背景温度当作零），我在准晶体生长中摹想一种类似情形。差别在于，这种最低能量的状态更加难寻找，而且原子“最佳”排列不能只靠一次加上一个原子，然后希望每个单独原子能解决自己的最小化问题就可以了。正相反，我们有一个全局性的问题要解决。解决方法是大量原子必须在同一时刻共同出力。我坚持认为，必须用量子力学方式才能得到这种合作；而进行合作的方式是，在线性叠加中原子的许多不同的组合排列同时一起“试验”（有点像在第9章结尾讨论的量子电脑）。最小化问题合适的解（虽然也许不是最好的）只有在单引力子（或别的适当的）水平达到时才能得到答案，而这只有当物理条件刚好适合时才会发生。

与头脑可塑性可能的关联

现在让我进一步推进这些猜测，并且询问它们是否和头脑的功能有任何关联。就我所见，这种关联最有可能呈现在头脑可塑性的现象中。我们记得头脑不完全像一台普通电脑，而是比较像一台持续不断改变的电脑。这些变化显然是树突棘成长或收缩导致该突触激发或退激发而引起的（见第9章第499页，图9.15）。我大胆猜测，这种成长或收缩可由类似准晶体成长的过程所制约。这样，可能不仅其他可能形态之一被试验，而且大量的复线性叠加形态也被试验。只要这选择的效应被维持在单引力子（或任何别的）水平之下，这些形态就能共存（而且按照量子力学的U步骤，几乎一定必须共存）。如果维持在这个水平之下，则可以开始进行同时发生的叠加计算，这和量子电脑的原则非常符合。然而，这些叠合似乎不太可能长期维持，因为神经信号产生的电场仍会严

重扰动附近的物质（虽然神经元髓鞘质的鞘有助于电绝缘）。我们猜测，这种计算的叠加实际上至少能维持一段时间，使得在达到单引力子（或别的什么）水平之前可以实际计算出某种有意义的结果。这类计算的成功结果会取代准晶体成长时简单的能量最小化的“目标”。这样，达到这个目标的过程正类似准晶体成功成长的过程！

这些猜测显然有许多模糊和疑虑之处。但是我相信它们之间的类比确实是有道理的。晶体或准晶体的成长受到它附近适当原子和分子浓度的严重影响。人们可以同样摹想，树突棘族的成长或收缩会同样受到周围的神经传递物质浓度的影响（譬如可能受到情绪的影响）。任何原子排列最终化解（或“缩减”）成准晶体的实在性都涉及能量最小化问题的解答。所以我以类似的方式猜测，在头脑中浮现的实在思维又是某问题的解答，但是现在这问题不只是能量最小化而已。它会涉及性质更复杂得多的目标，即有关头脑的计算方面和能力的需求和意图。我猜想，从原本是线性叠加的不同选择中寻找解答和意识思维的作用有密切的关系。这所有都与制约在U和R界限之间的未知物理有关。我现在宣称，这些未知物理的答案有赖于还未发现的量子引力理论—CQG！

这样的一个物理作用在性质上会是非算法的吗？我们回想一下第4章内描述的一般镶嵌问题。人们可以摹想原子组合问题会具有这种非算法的性质。如果这些问题能用我暗示的手段“解决”，则在我设想的头脑行为类型中的确有非算法因素的可能性。然而由此推理，在CQG中必须有某些非算法的因素。这里显然有许多猜测成分。但是依我看来，按照上面的观点，一定需要具有非算法特征的某些东西。

这类头脑联结的变化能发生多快？这个问题在神经生理学家之间有点争议。但是，由于永久的记忆可在十分之几秒的时间内记录下来，所以有关的改变可在这种时间内实现是不无道理的。为了使我自己的观点有成功的机会，这类速度确实必要。

意识的时间延迟

我想接着描述两个对人类实行的实验（在Harth　1982的书中所描述的），这些发现对我们这里的考虑具有惊人的含义。这些和意识的主动与被动行为所需的时间有关。第一个实验有关意识的主动作用，而第二

个是被动作用。合在一起，则含义会更加显著。

第一个实验是H.H.康胡贝和他的助手于1976年在德国进行的（*Deeke, Grötzinger, and Kornhuber 1976* ）。一些病人自愿把他们头上某一点的电信号（用人脑电流计，即EEG）记录下来，他们被要求在不同时刻完全出于自己的选择把自己右手的食指突然弯曲。其想法是，EEG的记录可以表示在脑壳内发生的某种精神活动，即参与弯曲手指的实际意识决定。为了从EEG追踪中得到有意义的信号，必须把几个不同的追踪试验平均一下，得到的信号不是很确定。然而，人们发现了一种很令人注意的现象，在手指实际弯曲之前整整一秒钟，或许甚至一秒半，从记录可以看到电位逐步在上升。这似乎表明，意识的决定过程需要超过一秒钟时间才会有行动出现！这种情形可以和另一情形相对照，当反应模式预先设定时，对外界信号产生反应所需的时间要短得多。例如，手指的弯曲不是出于“自由意志”，而是对闪光信号产生反应。在这种情形下，大约为1/5秒的反射时间是正常的，这大约比在康胡贝数据中（图10.5）检验的“自愿”行为快5倍左右。

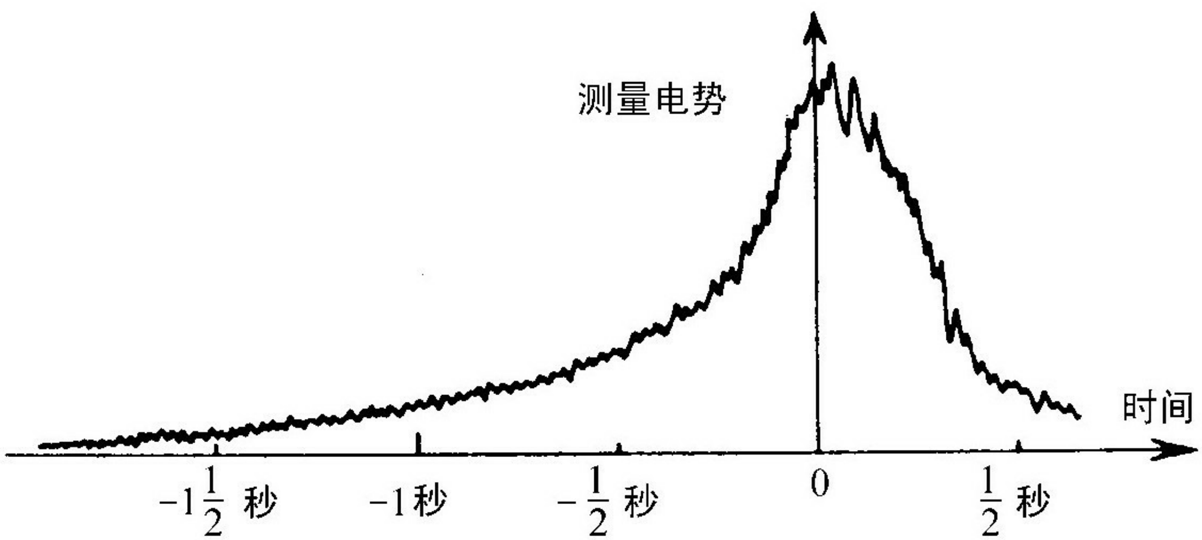


图10.5　康胡贝实验。在时间0作出弯曲手指的决定，而预兆的信号（基于许多试验的平均）暗示试图弯曲的“先知”

在第二个实验中，加州大学的本杰明·李伯特和旧金山锡安山神经研究所的贝特朗姆·费恩斯坦合作（Libet et al.1979），检测了必须进行脑外科手术的病人（进行手术的原因与该实验无关），并且同意把电极放在他们头脑触觉皮质的点上。李伯特实验的结果是，当刺激作用于这

些病人的皮肤时，他们大约需要半秒钟才能知觉到刺激，尽管头脑本身只需要大约1/100秒的时间接收到这个刺激信号，而且头脑能在大约1/10秒内得到预编程序的对这种刺激的“反射”反应（参阅上文和图10.6）。此外，尽管在刺激到达知觉之前有半秒钟的延迟，病人本身会有主观印象，以为在他们知觉到刺激时根本就没有发生过延迟！（李伯特的一些实验涉及丘脑的刺激，参阅478页，其结果类似触觉皮质的刺激。）

我们记得，触觉皮质是大脑中感觉信号进入的区域。所以，在触觉皮质上对应于皮肤某个特殊点的电刺激会使病人觉得犹如某种东西实际上触及皮肤上那一点。然而人们发现，如果该电刺激过于短促（短于半秒钟），则病人根本没有任何感觉。这情形可以用来和直接刺激皮肤上某一点的反应相对照，因为皮肤上一瞬间的接触都能被感觉到。

现在假定皮肤首先被触及，然后在触觉皮质的对应点加上电刺激。病人感觉到了什么？如果电刺激是在接触皮肤之后的1/4秒左右开始，则根本不会感觉到皮肤接触！这种效应被称为往前遮盖。刺激皮质在某方面用于防止正常皮肤接触的感觉被有意识地感觉到。只要事件发生在知觉感觉之后大约半秒钟之内，它就会被这后面的事件所阻止（“遮盖”）。这作用本身告诉我们，这一种感觉的知觉意识是在产生该感觉的实际事件后大约半秒钟发生！

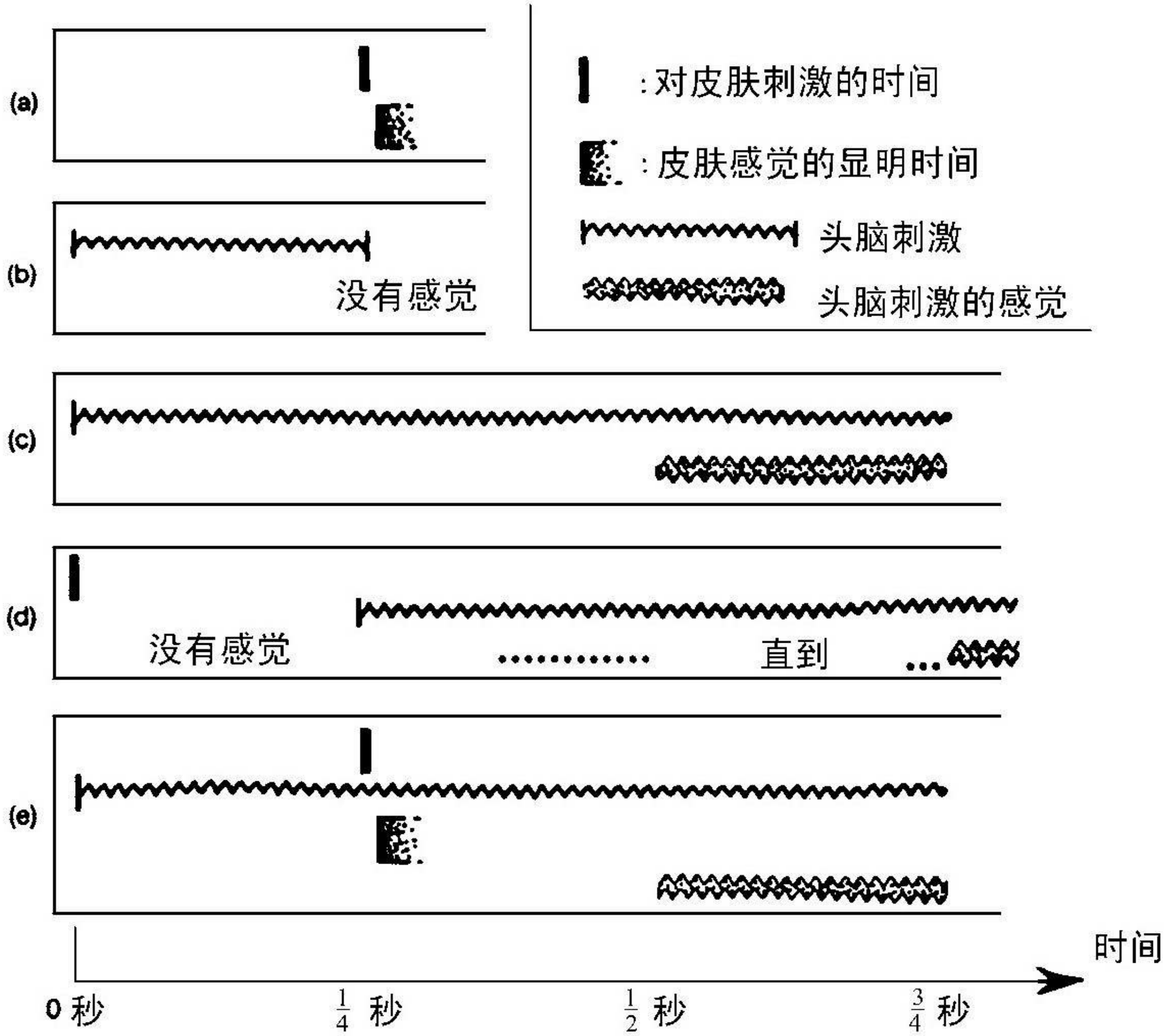


图10.6　李伯特实验（a）对皮肤的刺激“似乎”在接近该刺激实际时刻被知觉。（b）比半秒短的皮质刺激未被知觉。（c）比半秒长的皮质刺激在半秒后被知觉。（d）这样一个皮质刺激能够“往前遮盖”早先的一个皮肤刺激，这表明皮肤刺激的知觉实际上直到皮质刺激的时刻还没发生。（e）如果在这种皮质刺激之后很短的时间内加上一个皮肤刺激，则皮肤知觉被“认为属于以前的”，但是皮质知觉并非如此。

然而，人们似乎并没“感觉到”知觉延迟了这么长久的时间。赋予这个奇怪发现的意义的一个方法是，想象人所有“知觉”的“时间”实际上是从“实际时间”延迟大约半秒钟，犹如人们内部的钟“错”半秒钟左右。一个人感觉事件发生的时间总是在该事件实际发生的半秒钟之后。这就呈现出一幅协调的感觉印象的图像，虽然存在令人困惑的延迟。

也许在李伯特实验的第二部分可以证实这类性质。他首先对皮质进行电刺激，这个刺激延续比半秒还要长久许多的时间，一面进行刺激同时并接触皮肤，不过要从电刺激开始后半秒之内开始接触。不管是皮质刺激还是皮肤接触都被分别感觉到，而且病人很清楚分辨出两个刺激。当询问哪一个刺激先发生，病人却会说皮肤接触在先，尽管在事实上皮质刺激先开始！这样，病人看来把皮肤接触的知觉在时间上大约往回倒退半秒钟（图10.6）。然而，这似乎不是内部知觉时间的整体“错误”，而仅是在感觉事件时序上更微妙的重新安排。对于皮质刺激，假定在开始刺激后半秒之内被实际知觉到，则似乎不会以这种方式向过去回溯。

我们从上述的第一个实验可以推导出，意识行为在它发挥作用之前需要1秒或1秒半的时间，而根据第二个实验，似乎要在外界事件发生了半秒钟之后才意识到该事件的发生。想象一个人对某个未预料到的外界事件反应时会如何。假设这反应需要瞬间意识思考。根据李伯特的发现，意识作用之前必须花费半秒的时间；而且然后如康胡贝的数据所隐含的，必须花费比一秒还多很多的时间，人们“意志”的反应才能进行。从感觉输入直到动作输出的整个过程需要2秒钟左右！把这两个实验放在一起的明显含义是，如果反应是在大约2秒钟之内产生，则意识根本未起作用！

时间在意识知觉中的奇怪作用

我能够完全相信这些实验吗？如果这样的话，我们便被迫接受这个结论：当我们需要用少于1秒或2秒的时间采取行动去修正一个反应时，我们的行为完全像一台“自动机”。比较神经系统中的其他构造，意识无疑是行动迟缓的。我本人注意到这类事件，譬如正在我用手关车门的瞬间，无可奈何地看到在车子里还有一样东西待取出来。而我要停止手动作的意志命令进行得奇慢，以至于来不及阻止关门。但这真需要整整1秒或2秒的时间吗？我觉得不像要用这么长的时间尺度。当然，我对车中物体的知觉，加上我想象的“自由意志”命令去阻止我的手，我的有意识知觉都可以在这两个事件之后才发生。也许意识仅是旁观者而只不过经验到这出戏的“重演”而已。相似地，从表面上看，根据上面的发现，譬如当一个人在打网球时，不会有时间让意识起任何作用，在打乒乓球时更是如此！无疑这些球类的专家用小脑控制为他们所有的主要反应预先编好极佳的程序。但是，若说意识对于何时应打何种球路没有任何作

用，我有点难以同意。无疑必须预测对手将会做什么，而且对于对手可能的每一招都准备好许多预编的程序来反应，但我觉得这不很有效率，而且我难以接受在这期间一点也没涉及意识。这类评论对于日常交谈更为恰当。还有在交谈中，虽然人们有点能预期别人会说什么，但在别人的评论中总是经常发现一些出乎意料的东西，否则交谈就变得完全不必要！在通常的交谈方式中肯定不必花2秒钟那么长的时间去对别人反应。

也许有理由怀疑康胡贝实验足以证明意识“实际”需要1秒半时间来行动。虽然弯曲手指的意图的所有EEG追踪平均早那么许多就出现信号，也许只有某些情形这么早就有弯曲手指的意图，而这个有意识的意图也许实际上没有实现，并且有其他许多情形意识行为产生时刻距离手指弯曲的时刻要接近得多。（的确，后来一些实验结果导致和康胡贝不同的解释，参阅*Libet 1987* ，1989。尽管如此，我们仍对意识定时的问题深感困惑。）

我们此刻先认为两个实验结论实际上成立。我将做一个与此相关而令人忧心的设想。我认为，当我们考虑意识时对时间使用通常的物理规则，可能实际上犯了极大的错误！的确，时间实际进入我们意识感觉的方式有种非常古怪的性质。我想，若我们试图把意识知觉放进传统时序框架中，则可能需要非常不同的概念。意识毕竟是一种我们知道的现象，根据这种现象时间必须“流逝”！现代物理学处理时间的方式和处理空间 [71] 没有什么根本的不同，而物理描述的“时间”根本没有真正“流动”；我们只有一个显得静止的固定的“空间——时间”，在空间——时间的框架里展开我们宇宙中的事件！然而，根据我们的知觉，时间的确在流动（见第7章）。我的猜想是，这里也有些幻觉，我们知觉的时间不是“真的”完全像我们感觉到的以线性流动方式向前流动（不管这个含义是什么！）。我宣称，我们“表面”感觉到的时序是我们强加在感觉上的，以便理解我们的感觉和外在物理实在的均匀前进的时间之间相关联。

许多人也许在上面论述中找到大量哲学上的“不坚实之处”，他们这种指责无疑是正确的。一个人怎么可能对自己实际感觉的东西“弄错”呢？是的，按照定义，一个人实际知觉正是他直接发觉的东西，所以他不会弄错。尽管如此，我以为我们对于时间进展的知觉确实很可能是错的（虽然我无法充分使用平常语言去描述这信念），而且存在一些

支持这些信念的证据（见*Church Land 1984*）。

莫扎特“一瞥”即能捉住整篇“即使很长”的乐谱就是一个极端的例子（532页）。从莫扎特的描述中，人们必须假想这“一瞥”包含了整个乐曲的精华。然而，用通常的物理术语，这个知觉意识行为的这段实际外在时间，根本无法和表演乐曲所需的时间相比较。人们也许想象，莫扎特的知觉会采用完全不同的形式，也许像视觉景观或像一下子写出整篇乐谱一样用空间分布的方式。但是，即便是音符也需要相当长时间去精读，所以我非常怀疑莫扎特最开始用这种方式来知觉他的乐曲（而他一定会这么说！）。视觉景观似乎更接近他的描述，但是（以我本人较熟悉的，最普通的数学景象）我极其怀疑会有任何方法可以把音乐直接译成视觉语言。依我看来，更有可能的是，莫扎特“一瞥”的最佳解释必须纯粹是音乐性的，但和聆听（或表演）一段音乐有不同的时间内涵。音乐是由需要一定时间去表演的声音组成，这种时间在莫扎特的实际描述中允许“……我的想象使我听到了它。”

请听J.S.巴赫的“赋格曲艺术”四重赋格曲的最后一部曲。所有能体会巴赫音乐的人在这乐曲演奏10分钟，刚进入第三主曲音乐休止的时候没有不被感动的。整个曲子仿佛还在“那里”，但是现在一眨眼工夫就从我们耳边渐渐消隐而去。巴赫在完成这一个作品前死去，而他的乐曲就停在那一点，他没有留下任何只言片语表明他想如何继续。然而，该作品从开头就充满着自信和熟练，不能想象在那时刻他的头脑中没有完全掌握整个曲子的精华。当他尝试种种改善时，他是否需要在脑袋里以正常演奏节拍一次次从头到尾对自己演奏和尝试呢？我不能想象他如此进行。和莫扎特一样，他必须是把作品及其赋格曲整个孕育出来，以乐章所必需的复杂性和艺术性全部一起涌现出来。然而，音乐的时间品格是它的一种基本要素。如果不在“真实时间”里表演，它还成音乐吗？

小说或历史的产生可以呈现可互相比较（虽然困惑较少）的问题。在了解某人的一生时，必须思考他一生中不同的事件，好像必须在“真实时间”内重演才能评价这些事件。然而这并不必要。实际上，人的回忆能把以往费时的经验仿佛“压缩”到一瞬间内，并把它“重新经过”了一遍！

在音乐作曲和数学思考之间或许有一些强烈的相似性。人们也许认为数学证明要通过逻辑过程来获得，每一步都跟着前面那一步。然而，

产生新论证概念全不以这种方式进展。在建立数学论证中必须有全局和似乎模糊的概念内容；它和完全理解依序证明所需的时间没有什么关系。

假设我们接受意识的时间与时间前进不符合外在物理现实，我们不就面临着佯谬的危险吗？假定甚至意识效应具有一种模糊而目的论的性质，使得未来的印象就可能影响过去的行为。是的，这会把我们的导向矛盾，正如我们在第5章结尾考虑过的超光速信号佯谬的含义一样（参阅273页）。我们已经正确地排除掉这个佯谬。我想提议，正由于我所主张的意识实际所获得的性质，不必要存在佯谬。回顾一下我的设想，在本质上意识“看见”了某些必要的真理；而且它可以代表和柏拉图理想数学观念世界的某种接触。我们记得，柏拉图世界本身是没有时间的。知觉到柏拉图真理不携带有真正的信息，“信息”的技术意义是指信号的传递。因此，即使意识知觉在反时间方向传递也不存在实际冲突！

但是，即使我们接受意识本身和时间这种奇怪的关系，在某种意义上，它代表外在物理世界和没有时间性世界之间的接触，这些怎么可能和物质头脑的生理决定和时序行为相一致呢？如果我们不想扰乱物理定律的正常进行，我们的意识仅仅剩下一个纯粹“观察者”的角色。然而，我论断意识具有主动而且的确有力的作用，并且具有强大选择优势。我相信，这个难题的答案取之于CQG在解决两个量子力学过程U和R之间的冲突时，CQG的行为必须采用的奇怪方式（参阅447页，464页）。

回忆一下，当我们试使过程R和相对论（狭义）相协调时所遇到的时间问题（第6章367页，第8章470页）。按通常的空间——时间条件来描述这个过程似乎没有任何意义。考虑一对粒子的量子态。通常这样一种状态会是一个相关态（也就是说，不具有简单的|Ψ〉|〉形式，这里|Ψ〉和|〉各自描述一个粒子，而具有像|Ψ〉|〉+|α〉|β〉+……+|ρ〉|σ〉总和的形式）。那么对其中一个粒子进行观察就会以非局部的方式影响另一个粒子，它不能按照和狭义相对论一致的通常的空间——时间概念来描述（EPR；爱因斯坦——波多尔斯基——罗森效应）。这种非局部效应会隐含涉及我提议过的树突棘成长和收缩的“准晶体”相似性。

我在这里以如下意义来解释“观察”，即把每一个被观察粒子的作用一直放大达到类似CQG的“单引力子”水平。如果应用更“传统”的说法，则“观察”更加模糊。当一个人必须认定自己的头脑一直在“观察头脑自

身”时，很难看出他怎么能着手发展头脑行为的量子理论描述！

我本人的想法是，在另一方面，CQG提供了一个态矢量缩减（R）不必依赖任何意识思想的客观物理理论。我们还未得到这个理论，但是至少寻求此理论不会受到“意识到底‘是’何物”这一深奥的问题的阻碍！

我想象，一旦真正寻求到CQG，那时就可能依照它来揭示意识的现象。我相信当得到CQG时，该理论的必要性质离开传统空间——时间之描述将比离开上述令人困惑的两粒子EPR现象更远。正如我提议的，如果意识现象依赖这种想象的CQG，则意识本身用我们现在传统的空间——时间来描述将显得非常不适合！

结论：孩子的观点

我在本书中提供了许多议论，试图显示以下观点不能成立：我们的思维基本上和一台非常复杂的电脑的行为一样。这种观点在现代哲学探究中甚有影响。在人们明确假设执行算法本身就能唤起意识知觉时，采用了西尔勒“强人工智能”术语。有时以不太明确的方式采用诸如“作用主义”等其他术语。

有些读者可能从一开始即把“强人工智能支持者”当成稻草人！仅靠计算不能唤起快乐或痛楚；它也不能理解诗歌、夜空的美或者声音的魔力；它不能希望、恋爱或沮丧；它也不能具有一个真正自发的目的，这一切难道不是“显而易见”的吗？然而科学似要逼迫我们去接受，我们所有人仅是由非常精细的数学定律巨细无漏制约（甚至最终也许只是随机地）的世界中很小的一部分。控制我们行为的头脑本身似乎也由同样的精密定律所制约的。所呈现的世界图像是：所有这些精确的物理活动，实际上只不过是一个庞大的计算过程（也许是概率的），所以我们的头脑和精神只能按照这种计算来理解。也许当计算复杂得非同寻常时，它们便开始获得我们与“精神”这术语联想在一起的更诗意或更主观的品质。然而，在这样的图像中，非常难免总有缺少了什么的不愉快感觉。

我在自己的论证中试图支持以下观点，即任何纯粹计算的图像的确缺少了某些要素。然而我同时保持希望，将来通过发展科学和数学，在理解精神方面最终会取得根本的进步。这明显是个难题，但是我已试图指出存在一条真正的出路。可计算性和数学的精确性根本不相同。在精

确的柏拉图数学世界中，具有人们要多少即有多少的神秘和美，而大部分神秘从一些概念而来，这些概念属于柏拉图数学世界中较有限制的算法和计算以外的部分。

对我来说，意识是如此重要的现象，我简直不能相信它只不过是从复杂的计算“意外”得来的。宇宙的存在正是由于意识现象才被得知。人们可以争论道，若宇宙被不允许意识存在的定律所制约，就根本不是宇宙。我甚至愿意说，迄今为止所有人们提出的宇宙数学描述都不能达到允许意识存在的要求。只有意识现象才能把一个想象的“理论”宇宙变成真正的存在！

我在这些章节中提出的论证也许是过于曲折复杂。我承认有一些是猜测性的。同时我相信，有些是不可避免的，然而，在所有这些技术细节下面隐藏着一个感觉，即意识精神不能像一台电脑那样运行是“显而易见”的，即使真正涉及的精神活动中，有许多的确像电脑一样运行。

这种明显性连小孩都能看出来，虽然那个孩子在他后来的生命中被细心的推理和巧妙的定义选择所威吓，以至变成相信一些明显的问题“不是问题”。有时孩子容易看清楚的事情长大后却变得非常模糊。当“实在世界”的事务开始落到我们的肩膀上来时，我们经常忘记了孩提时代的惊奇印象。儿童们不害怕提那些使我们大人羞于启齿的基本问题：在我们死后每一个人的意识流会发生什么；在我们出生之前又在何处；我们过去曾经是或者将来会变成另外一个人吗；为什么我们会知觉；为什么我们在这里；为什么存在一个我们实际上居住的宇宙？这是些令人困惑的谜题，它随着我们每个人的知性觉醒而来，而且无疑也随着任何最早的生物或其他个体的真正自我知性觉醒而来。

我记得自己在孩提时代曾为许多这类困惑所烦恼。也许我的意识突然和别人的相互交换。假定每一个人都只带有属于他个人的记忆，我怎么能知道这种互换早先不曾发生在我身上？我如何向其他人解释“相互交换”的经验？它真有任何意义吗？也许我只不过不断重复生活在同样的10分钟经验中，每一次都具有完全同样的知觉。也许对我来说，只有现在时刻是“存在的”。也许明天或昨天的“我”实际上是具有独立意识而完全不同的人。也许我实际上活在时间往回走的情况下，我的意识流是朝向过去，因此我的记忆告诉我将要发生什么，而不是已经发生了什么。这样，学校里的不快乐经验实际上在等着我，而我很快就会非常不

幸地遭遇到。这个时间倒向和正常经历的时间进展之间的区别是否真正“意味”着什么，使得一个是“错”的，而另一个是“对”的？为了使这些问题在原则上得以解决，必须有一种意识理论。但是，甚至怎么可能开始向本身都不是有意识的实体去解释这类问题的内涵呢……？

注释

第1章

[1] 例如，参阅*Gardner 1958* ，Gregory 1981以及所引用的参考文献。

[2] 例如，参阅*Resnikof and Wells 1984* ，181—184页。有关计算奇才的经典总结参阅*Rouse Ball 1892* 以及*Smith 1983* 。

[3] 参阅*Gregory 1981* 285—287页，*Grey Walter 1953* 。

[4] 这个例子引用自*Delbrück 1986* 。

[5] 参阅*O'Connell 1988* 和*Keene 1988* 。参阅列维更多的有关电脑下棋的情形。

[6] 当然，大部分弈棋问题都被设计得使人类很难解决。要去构造一种人类觉得不是极难，而现代解决下棋问题的电脑在1000年内也解不出的下棋问题似乎不太困难。（所需要的是一个每下一着都要筹划非常多步的，但又是相当明显的方案。例如，已经知道一些需要筹划大约200步就绰绰有余的问题！）这提出了一个有趣的挑战。

[7] 为了明确起见，我在本书从头到尾地采用了西尔勒的术语“强AI”以表示这一极端观点。术语“机能主义”也被经常地用于表示本质上同样的观点，但也许不总是这么明确。明斯基（1968），伏多（1983）以及霍弗斯达特（1979）是这类观点的一些倡导者。

[8] 可从*Searle 1987* 211页找到这种宣称的一个例子。

[9] 道格拉斯·霍弗斯达特在对西尔勒的复印在《精神》上的原始论文的批评中抱怨道，由于涉及的复杂性，没有一个人可想象把另外的一个人脑的整个描述“内在化”。的确不能！但是就我所见，这不是问题的全部。人们仅仅关心实行目前要体现单个精神事件发生的一段算法的那

个部分。这可以是在回答图灵检验问题时某个瞬息的“意识实现”，或者它可以是某种更简单的东西。任何这种“事件”是否都需要一段极其复杂的算法呢？

[10] 参见载于*Hofstadrer and Dennett 1981* 368页和372页的*Searle1980* 的论文。

[11] 有些关于这类事体博学的读者也许会忧虑符号不同的问题。但是，如果我们在进行交换的同时，使其中的一个电子旋转360°，甚至那个（可争论的）区别也消失了！（参见第6章357页的解释。）

[12] 参见*Hofstadrer and Dennett 1981* 的导言。

第2章

[1] 我是采用当代通常的术语，它现在把“0”包括在“自然数”之中。

[2] 还有许多把一对，3个等数编码成单独一个数的其他方法。虽然它们对于我们现在的目的不甚方便，数学家却熟知这些方法。例如，公式$1/2[(a+b)^2+3a+b]$便是用一个单独的自然数来代表一对自然数（a, b）。试试看！

[3] 我在上面没花工夫去引进某种表示起始一个数（或指令等）的序列的记号。这对于输入没有必要，由于当遭遇到第一个1时事情刚刚开始。然而，对于输出需要某些其他东西，这是由于人们预先为了达到第一个（也就是最左边的）1不知道要沿着输出磁带看多远。尽管在往左看时会遇到0的很长的串，这并不能保证在左边更远处不再有1。人们对此可采用不同的观点。其中一种总是用特殊记号（譬如，在收缩步骤中用6来编码）去起始整个输出。但是为了简单起见，我在自己的描述中将采用不同的观点，也就是总“知道”仪器实际上已遭遇到了多长的“磁带”（例如，人们可以想象它留下了某种痕迹），在原则上不必去检查无限长的磁带，就能肯定整个输出已被查过。

[4] 一种把两盘磁带的信息编码到单独一盘磁带上的方法是插入法。这样，这盘单独磁带上的奇数号码的记号可代表第一盘磁带的记号，而偶数号码的记号可代表第二盘磁带的记号。可用类似的方案来处理3盘或更多盘磁带。这一过程的“低效率”起因于如下事实，即阅读机

必须沿着磁带不断地来回进退，并在上面留下记号以记住在该磁带偶数和奇数部分的什么地方。

[5] 这一过程只是指做过记号的磁带可解释作自然数的方法而言。它并不改变我们特定的图灵机的编号譬如EUC或XN+1。

[6] 如果T_n 没有被正确地指定，则U只要在n的二进制表示中到达多于4个1的第一串，就会像n的数已被终止那样进行下去。它就会把该表示式的余下部分当成m的磁带的部分来读，所以它会继续进行某种毫无意义的计算！如果需要的话，可采用扩展二进制记号来表示n，这种特征就能被清除。我决定不这样做，以免使这台可怜的通用机器U更加复杂！

[7] 我感谢大卫·德义奇根据以下我得到的u的二进制形式推导出十进制形式。我还感谢他检验u的这个二进制值实际上的确给出了一台通用图灵机！事实上u的二进制值为：

1000000001011101001101000100101010110100011010001010000
11010100110100010101001011010000110100010100101011010010
01110100101001001011101010001110101010010010101110101010
01101000101000101011010000011010010000010101101000100111
01001010000101011101001000111010010101000010111010010100
11010000100001110101000011101010000100100111010001010101
10101001010110100000110101010010110100100100011010000000
01101000000111010100101010101110100001001110100101010101
01010111010000101010111010000101000101110100010100110100
10000101001101001010010011010010001011010100010111010010
01010111010010100011101010010100100111010101011000110100
10101010111010100100010110101000010110101000100110101010
10100010110100101010010010110101001001011101010100101011
10101001010011010101000011101000100100101011101010100101
0111010101000000111010100100000110101010100101110101001010110100
10010010101110100000101011101000010001110100000101010011
10100001010011101000000100010111010001000011101000000100101
00111010001000010110100010100101110100010100101101001000
00101101000101010010011010001010101011101001000001110100

1001010101011101010101001101001000101011010010010010110
0000000101101000001000110100000100101101000000000110100
0100010111010010101000110100101001010110100000100111010
1010100101101001001110101000000101011101010000000110101
0001010101101001010101101010000101011101010010010101110
0100010010110101001000010111010000001110101001000101101
1001010011010101000101110101001010010111010101000001011
0101010000010111010000001110101010000101011101001010101
0101010000101110101000101010111010101001001011101010101
0001110101000000011101001001000011010010010001011010101
1010011101000000000101101001000011010101010100101110100
0001101001000101010111010000100011101000100001110100001
0100000001011010000010010111010101001010101101000100010
1011101000000100111010101001101000000101010110100001000
1010010000100011101010101010100111010000100100111010001
0100000111010000101001011010000101000011101010101010101
0100010010011010001001001101010010100101110100010001010
1101000000011101000100010010111010011010010010000101101
1010100110100010100010111010000110101000010001011010100
1010101001010010110101010011010010010101110100110100100
0010110100010101010001110100100001010110100000001001101
1000100101110100100001101010000010010111010010010100111
0100101010110100110100100101001011010011010010100000101
0100100000111010100100110101010100001011101001010000101
1010010101010111010100010010110100100111010010101000101
1010001001110101000010110100100111010010101010101110100
0001110100101010100101110100100011101010000010101011100
10101000001011010010011101010000001011101001011010100000101011
1011010100110101000100010110101010100101110101000101001
1101000101010101110100100001010110101000101110101001001
1010111010101001001011101010001110101000111010100100100
0111010100011101010010100010111010100010111010100001001
1110101000111010001010001011101001010010111010100101010
1011101001010101010101101010000101010101101000010011101
0001010101010111010101000101011101010100010101110100000
1110101010001001011101000000111010101001000101110101000

00110101000010110100000011101001000000101110101000111010
10010001010111010100110101010100010101101000001101010101
00101010110100000010011010101010010011101010011010101010
01001011010100110100100100111010000011010101010100101011
01010001001101000101001010101110100000110101010101010010
11010001000111010001010101010101101000100011101000010101
1101000100100001110100110100000000100111010000000100101110
100010001010011101000000010010111010010101010100101101000
01010101011101000100101001011101000001000101110101010010
110100010001001110100000010010101110100000010101011010000
1000111001111010000100000011101000010010011101000000101001
011101000000101001011010000100010101110100001000100110100
01000011101011110100001001001011101000010010010111010000
00010101110100001010100011010001001011101000010000011101
0000100111010001000000101110101010010110100010000001011101
00001010101011101000000010101011101000100001010111010001
0000101011101001000000111010100100100110100000001010111010
0010001001011101010100001110101001010110100101010100001
1010000001010011010000000111010000001001001110100101101001
00010100101101010100110100010100100101101010100110100010
10100010110011010100100101110101010011010001010101010110
01101010001010101100110100100010101010111010001000111010
01001010101010110100101001010001101001000000101110100000
1101010100101010101101001010101101001000100010111010001010101110
01110101001010101010111010010110100100100010101100110100
10010010101011101001101001001001010110100101101001001001
00101101001011010010010100010110011010010010100101011101
00010101110100100101110011010010010101001011100110100101
00010101011101000100011101000010100101101001010001011101
00101000101011010001001110100101000100101110100010011101
00101001000101110011010010001000111010001001110100101001
01010111001101001010000011100110101010101011010000000111
01001010100101010111010010001110100101010010101110011010
00010100100110011010100000110100000001110100101010100101
01110011010100010000110100000001110100010010101010111010
00100011101010101010101010110100001001110100100010010101

1101001010100010011010100000001011010010001110101000010101110100100001101010000000101101001000111010100100101110100001101010000101010110101000101110101000010100101110101000101110101000101010101110011010100010101101000011010100001001010

有进取心的读者可把一台效率高的家庭用电脑，以正文中给出的方法，应用于不同的简单的图灵机号码中，验证上面的号码事实上的确给出了一台通用图灵机的动作行为！

对图灵机的不同明记方法可使u的值降低一些。例如，我们可以免除STOP，而相反的采取这样的规则，即只要机器在某个其他非0的内态后重新进入内态0时它就停止。这样做没有太多收益（如果有的话）。如果我们允许磁带有比仅仅0和1的更多的记号，则能得到更大的收益。在文献中的确描述过显得非常简洁的通用图灵机。但是，由于它们一般地依赖于图灵机描述的极其复杂的编码，所以这种简洁性是骗人的。

[8] 参阅德伏林（1988）的和这一著名断言相关事体的非技术性讨论。

[9] 我们再次简单地应用前面进行的步骤，当然也能击败这个改善了的算法。然后我们可以用这新的知识去进一步改善我们的算法；但是我们又可将其击败等。这一迭代步骤所导致的这类考虑将在第4章（参阅第141页）和哥德尔定理联系起来讨论。

第3章

[1] 见*Mandelbrot 1986* 。我所选取的特殊的放大序列是取自Peitgen and Richter 1986。该书中有许多五彩缤纷的芒德布罗集的图画。进一步的图解可参见*Peitgen and Saupe 1988* 。

[2] 尽我所知，要求对任意实数总存在某种确定其n位数是什么的规则的观点是协调的，虽然不是传统的。尽管这样的一个规则可以是无效的，甚至在一预定的形式体系中根本不能定义（见第4章）。我希望它是协调的，因为这正是我最希望坚持的观点。

[3] 关于是谁第一个得到这个集，实际上存在一些争议（见*Brooks*

and Matelski 1981，Mandelbrot 1989 ）；但是这一争议本身的存在更加支持了这一集合是被发现而非发明的观点。

第4章

[1]　在考虑其元素又为集合的集合时，我们必须小心地区分该集的元以及该集的元的元。例如，假定S是另一确定的集T的非空子集的集合，而T的元素是一个苹果和一个橘子。T就有“二性”而非“三性”的性质；但是S实际上有“三性”的性质；S的3个元是：一个只包括一个苹果的集合，一个只包括一个橘子的集合以及包括一个苹果和一个橘子的集合——总共有3个集，这就是S的3个元素。类似地，其元素只有空集的集合具有“一性”而非“零性”——它有一个元，也就是空集！空集本身当然只有零个元。

[2]　事实上，哥德尔定理的推理可以用这种方式来表达，使之不依赖于诸如P_k　（k）的命题“真理”的完全外在的概念。然而，它仍然依赖于某些符号的实在“意义”的解释：尤其是，“～∃”的真正意义是“不存在（自然数）……使得……”。

[3]　在下面用小写字母代表自然数，而用大写字母代表自然数的有限集合。令m→[n,　k，r]表示陈述“如果X={0，1，……，m}，它的k个元素的每一子集都被放到r个盒子里，则存在X的一个“大”的至少包含n个元素的子集Y，使得所有Y的k元素子集都被放到同一盒里去”。这里“大”的意思是Y中的元素数目比作为Y中的元素的最小的自然数还大。考虑如下命题：“对于任意选取的k, r和n，存在一个m_0 ，使得所有大于m_0 的m，陈述m→[n, k，r]总成立。”J.巴黎斯和L.哈林顿（1977）指出这一命题等效于算术的标准的（皮阿诺）公理的哥德尔型命题。这一道命题是不能从那些公理证明得到的，但是关于那些公理作了某些“显然正确”的断言，正也就是，在这种情形下，从公理推断出来的命题本身是真的。

[4]　其题目为《基于序数的逻辑系统》，而且有些读者将会熟悉我用在下角标示代表康托尔序数的记号。使用我在上面所描述的步骤得到的逻辑系统的等级由可计算的序数来表征。

存在一些相当自然的和容易陈述的数学定理，如果人们试图用标准

算术的（皮阿诺）法则去证明，就需要使用在前面概述的“哥德尔化”步骤。这些定理的数学证明根本就不依赖于任何模糊或可疑的似乎处于正常数学论证的步骤以外那一类推理。参见斯莫林斯基（1983）。

[5] 在第3章111页提及的连续统假设（即$C=\aleph_1$ ）是我们在这里遇到过的最“极端的”的数学陈述（虽然人们还经常考虑比这更极端得多的陈述）。连续统假设，由于哥德尔本人和保罗—J.科恩确立了它实际上和集论的标准公理和步骤法则无关，而变得格外有趣。这样，对连续统假设的看法可用来区分形式主义者和柏拉图主义者。对于一个形式主义者而言，由于用标准的（策梅洛——弗兰克尔）形式系统既不能证明也不能否定连续统假设，所以是“不可判定的”，把它叫作“真”的或“假”的都“没有意义”。然而，对于一个好的柏拉图主义者，连续统假设或是真的或是假的，但为了确立它就需要某种新的推理形式——实际上甚至超出了对策梅洛——弗兰克尔形式系统使用哥德尔型命题的手段。[科恩（1966）本人提出一种使得连续统假设成为“显然错误”的反思原则！]

[6] 参阅*Rucker 1984* 的生动而不太专业性的有关论述。

[7] 布劳威尔本人似乎部分地因为对自己的一个拓扑学的“布劳威尔不动点定理”证明的“非构造性”忧虑，而开始沿着这个思路思考的。该定理断言，如果你取一个圆盘——也就是一个圆和它的内部——并且以一种连续的方式运动到它原先定位的区域的内部，那么在圆盘上至少有一叫作不动点的点，它刚好在自己开始的那点结束。人们也许不知该点准确地在何处，或者也许那里有许多这类的点，这个定理只是断言某一这类的点的存在。作为数学的存在定理而言，这实际上是一个相当“构造性的”定理。依赖于所谓的“选择公理”或“佐恩引理”的存在定理具有不同程度的非构造性（参阅*Cohen 1966*，*Rucher 1984* ）。布劳威尔情形的困难和下面相类似：如果f为一实变量的实数值的连续函数，该函数既取有负值又取有正值，找到f取零值的地方。其通常的步骤是涉及重复地对分f改变符号的区划。但是去决定f的中间值为正、负或零，在布劳威尔需要的意义上可以不是“构造性的”。

[8] 我们列出集合{v，w，x，……，z}，这里按照某种字典方案，v代表函数f。我们在每一阶段（递归地）检验看是否有f（w， x，……，z）=0，当仅当在这种情形下，才有命题∃w， x，……，z[f（w， x，……，z）=0]。

[9] 最近妮诺·伯龙（由于受这本书初版精装本中我的议论所刺激）告诉我，她已能断定芒德布罗集（的补集）在下面注释10的特殊意义下的确是非递归的，正如我在正文中所猜测的那样。

[10] 存在实数的实值函数的可计算性的一种新理论（和传统的自然数的自然数值函数相反），由伯龙、苏伯和斯马勒（1989）提出。我最近才注意到该理论的细节。该理论还适用于复函数，它可能和正文中提出的问题有重要的关系。这一新工作的一些结果赋予芒德布罗集在适当意义下为非递归的猜测以强大的支持。

[11] 这一特殊问题可更正确地被称为“半群的字问题”。还有其他形式的字问题，其法则略为不同，我们在此不予关注。

[12] 汉弗（1974）和迈尔斯（1974）还指出，存在一个单独的（一个巨大数目的花砖）集合，它只能以不可计算的方式来镶嵌平面。

[13] 事实上对于大的n，这一步骤的数目可用一些技巧减少到nlogn loglogn——这当然还在P中。参见克奴斯（1981）有关此类问题的更多资料。

[14] 更正确地讲，只对是/非类型问题（例如，给定a, b和c, a×b=c为真的吗？）P、NP和NP完全问题的族（参见187页）才被定义，但在正文中的描述对我们的目的已经足够。

[15] 严格地讲，我们需要是/非的模式，诸如：推销员是否有一条距离小于若干的路径呢？（见上面的注释14）。

第5章

[1] 一个显著的事实是，所有确立的和牛顿图像的偏差都以某种基本的方式和光的行为相关。首先，存在麦克斯韦理论中的脱离物体的携带能量的场。其次，正如我们就要看到的，光速在爱因斯坦狭义相对论中起着关键的作用。第三，只有当运动速度和光速可相比较时，爱因斯坦广义相对论和牛顿引力论的微小偏离才变得显著。（太阳引起的光偏折，水星运动，在黑洞中和光相比较的逃逸速度，等等。）第四，首先是在光的行为中观察到量子力学中存在的波粒二象性。最后，还有量子电动力学，它是带电粒子的量子场论。可以合情理地猜测，牛顿本人已

经准备接受他的图像躲藏在光的神秘行为后面的深刻问题。参阅Newton 1730；还可参阅Penrose 1987a）。

[2]　有一个美妙的、很成功的物理理解的实体——亦即卡诺、麦克斯韦、开尔文、玻尔兹曼等的热力学——我在分类时忽略了它。这可能引起某些读者的困惑，但我是故意这么做的。因为某种在第7章会变得更清楚的原因，我本人非常犹豫是否将其归于超等理论的范畴中。然而，许多物理学家也许会认为把这么漂亮的根本的观念放到低到仅仅有用的范畴是亵渎神物！依我看来，热力学按通常的理解是某种仅适合于平均的东西，而不适合于系统中的每一个别部分——部分的原因是由于它为其他理论的推论——在我这里的意义上不是一个完整的物理理论（这同时适用于作为其数学基础的统计力学的数学框架）。我以此事实作为借口以回避这一问题，把它们一块放到分类之外，正如我们在第7章将会看到的，我宣布在热力学和在前面已经提到的属于有用的范畴的大爆炸标准宇宙模型之间存在一种紧密的联系。我相信，在这两组观念之间（现在还缺一部分）的某种联合，在必需的意义上，甚至会被认为属于超等的范畴的物理理论。这是我还要在后面论述的内容。

[3]　我的同事们问我应将“扭量理论”归于何类。这是一种观念和技巧的精心集合，我自己曾为此花费了许许多多的心血。就扭量理论作为物理世界的一个不同理论而言，它只能被收到尝试的范畴中。

[4] 然而，伽利略经常用水钟来为其观察定时，见Barbour 1989。

[5]　用牛顿的名字来命名这个模型——的确就“牛顿”力学总体而言——仅仅是一个方便的标志。牛顿本人对于物理世界实际性质的观点似乎不像这么独断，而是更微妙，更难以捉摸[最有力地促进这一“牛顿”模型的人要算R.C.玻斯科维奇（1711—1787）]。

[6]　拉飞逸·索金曾向我指出，存在一种意义，在这种意义上，可用一种和对（譬如讲）牛顿系统所用的相似的方式来“计算”此一特殊的玩具模型。我们可摹想一个计算序列C_1，C_2，C_3，……，这些步骤允许将系统的行为计算到越来越后而没有时间的极限，并且不断增加精确性（参阅224页、225页）。在现在情况下，为了达到这个目的，我们可允许将图灵机动作T_u（m）进行N步定义为C_N，如果这一动作那时还不停止则“认为”T_U　（m）=□。然而，在T_u　（m）=□的地方，由引进牵涉到

诸如“对所有的qT（q）停止”的双重量化的陈述的演化，不难修正我们的玩具以战胜这类“计算”。（存在无限多对相差2的素数的未解决问题即为这种陈述的一个例子。）

[7] 正如第4章（575页注释9）提示的，新的伯龙——沙柏——斯马勒（1989）理论可提供一种在数学上更能接受的方法来解决其中的一些问题。

[8] 伟大的意大利/法国数学家约瑟夫·L.拉格朗日（1736—1813）比哈密顿早24年左右就知道了哈密顿方程。他虽然和哈密顿观点不一样。更早时期的一个同等重要的发展是力学中欧拉——拉格朗日方程的表达形式。这样牛顿定律可认为是从一个更高的原则，即最小作用量原理（P.L.M.德毛帕裘斯）推导而来。除了其伟大的理论意义之外，欧拉——拉格朗日方程还提供了具有显著威力和实用价值的计算步骤。

[9] 刘维尔相空间体积只是整族具有不同维数的在哈密顿演化下保持不变的“体积”（称作庞加莱不变量）之一。但是，我的这个断言如此之囊括无遗实在有些过分。我们可以想象一个系统，将其中我们不感兴趣的一些自由度（对某些相空间体积有贡献）“倾倒”到某处去（诸如逃到无限远处的辐射），这样我们感兴趣的部分的相空间就会减小。

[10] 这第二个事实尤其是科学的极大的幸运。因为没有它的话，巨大物体的动力行为就不可理解，而大物体行为几乎不能给精确适用于粒子本身的定律提供任何暗示。我猜想，牛顿之所以那么强调他的第三定律的原因在于，如果没有它，则从微观到宏观的动力行为的传递就不成立。

另一个对于科学发展生死攸关的“奇迹般的”事实是，平方反比律是仅有使围绕着中心物体的一般轨道具有简单的几何形状的方次律（随距离而减小）。如果定律或力是倒数律的或立方反比律的，开普勒还会有何成就呢？

[11] 我已为各种场选好了单位，以使和麦克斯韦原始方程的形式相接近（除了他的电荷在我处为$c^{-2}\ \rho$以外）。当用其他的单位制时，因子c的分布将会不同。

[12] 事实上，我们具有无限多的x_i 和p_i 。更复杂之处在于，我们不能只用这些场的值作为坐标，必须引进某种麦克斯韦场的“势”才能纳入哈密顿理论的框架中去。

[13] 也就是说，不是二次可微的。

[14] 洛伦兹方程告诉我们，由电荷所处的地方的电磁场引起了作用于它上面的力；如果它的质量又是已知的，牛顿第二定律就告诉我们该粒子的加速度。然而，带电粒子经常以近于光速的速度运动，狭义相对论的效应变得很重要，影响了实际上应取的粒子质量数值（见下一节）。正是这种原因使作用在带电粒子上的正确的力定律推迟到狭义相对论的诞生才被发现。

[15] 事实上，自然界中的任何量子粒子，在某种意义上，整个自身都像一台这样的钟。正如第6章要讲到的，任何量子粒子都和一个振动相关，其频率与质量成正比；见292页。现代最精确的钟表（原子钟、核子钟）归根到底是依赖于这个效应。

[16] 也许读者会忧虑，由于旅行者世界线在B处出现了一个“角”，正如图示的，他在事件B处遭受到无限大的加速度。可以用有限的加速度将他的世界线在B处的尖角弄圆滑，这只不过把他所经历的由整个世界线的闵可夫斯基“长度”所测量的总时间稍微改变一点。

[17] 这些就是M依照爱因斯坦的同时性定义，由从M发出并被问题中的事件反射回到M的光信号判断的事件空间。例如，见*Rindler 1982*。

[18] 这是该形状的初始的对时间的二阶微分（或“加速度”）。形状的改变率（或“速度”）在初始时为零，因为球面在开始时刻是静止的。

[19] 杰出的法国数学家埃利·卡当（1923）首先对牛顿理论的数学形式重新进行表述——这当然是在爱因斯坦的广义相对论之后。

[20] 在这种意义上的局部的欧几里得性的弯曲空间称作以伟大的贝纳德·黎曼（1826—1866）命名的黎曼流形。他在高斯的某些更早的有关的两维情形的工作之后，首先研究这类空间。我们在此需对黎曼观念作重大的修正，也即允许几何为局部闵可夫斯基的，而不是欧几里得

的。通常将这种空间称为洛伦兹流形（属于所谓的伪黎曼或更不逻辑点的半黎曼流形的一类）。

[21] 或许读者会忧虑，零值何以代表“长度”的最大值！事实上的确如此，只不过在空洞的意义上而言：零长的测地线的特征是，没有任何其他的粒子的世界线可将其上面的任何一对点（局部地）连接。

[22] 畸变效应和体积改变的分解事实上不像我所表达的那么明确。里奇张量本身会引起一定量的潮汐畸变。（对于光线而言，这种分解是完全明确的；参阅*Penrose and Rindler 1986* ，第7章。）例如可参阅*Penrose and Rindler 1984* 240页和210页关于外尔和里奇张量的精确定义。（德国出生的赫曼·外尔是20世纪一位杰出的数学人物；意大利的格里高里·里奇是一位有巨大影响的几何学家，他在19世纪奠定了张量理论。）

[23] 大卫·希尔伯特在1915年11月发现了实际方程的正确形式，但该理论的物理观念则完全归功于爱因斯坦。

[24] 对于通晓这些东西的读者而言，这些微分方程正是用爱因斯坦方程代入到完整的比安基等式而得到的。

[25] 存在某些（虽然不是非常令人满意的）方法可绕过这个论证，参阅*Wheeler and Feynman 1945* 。

[26] 因为它不是二维的，而是三维的，所以在这里用术语“超面”比“曲面”在技术上更为合适。

[27] 有关这些问题的严格定理一定是非常有用的，并且非常有趣，可惜迄今还没有得到。

[28] 现在这个理论是不可计算的，它的（临时的）毫无用处的答案是无限大。

第6章

[1] 我理所当然地认为，“严肃的”哲学观点应该至少包含足够分量的现实主义。当我得知一些显然严肃的思想家，经常关心量子力学含义

的物理学家采取强烈的主观观点，说在“那里”实际根本没有实在的世界时，总是十分吃惊！我尽量采用现实主义观点的事实，并不意味着我不了解某些人经常认真地坚持的这种主观观点，只是因为我认为它们没有意义。参见*Gardner 1983* 第1章对这种主观主义的强烈而风趣的攻击。

[2] 尤其是J.J.巴尔末在1885年注意到，氢光谱线的频率具有R（n^{-2} - m^{-2} ）的形式，其中n和m为正整数（R为常数）。

[3]　也许我们不应该太轻易地抛弃这种“全场”图像。爱因斯坦本人（正如我们将要看到的）彻底了解量子粒子呈现的分离性，耗费了最后的三十年去寻求对这一般经典类型的完全广泛的理论。但是，正和其他人一样，爱因斯坦的企图没有成功。除了经典场以外需要某些东西用以解释粒子的分离性。

[4] 杰出的美籍匈牙利数学家约翰·冯·诺依曼（1955）在他的经典著作中描述了这两种演化的过程。我把他的“过程1”叫作R——“态矢量的减缩”——他的“过程2”叫作U——“幺正演化”（这实际上表明概率幅在演化中守恒）。实际上，还有量子态演化U的其他（虽然是等效）的描述，人们在这种描述中可以不使用“薛定谔方程”。例如，在“海森伯图像”中，态被描写成根本不演化，而动力学演化被归结为位置/动量坐标意义的连续移动。这些差异在这里对我们不重要，过程U的不同描述是完全等效的。

[5]　为了完整起见，我们必须列举出所有需要的代数定律。按照在正文中使用的（狄拉克）记号，它们可写成：

$|\Psi\rangle + |x\rangle = |x\rangle + |\Psi\rangle$ ，

$(z+\omega)|v\rangle = z|\Psi\rangle + \omega|\Psi\rangle$ ，

$z(\omega|\Psi\rangle) = (z\omega)|\Psi\rangle$ ，

$|v\rangle + 0 = |\Psi\rangle$ ，

$|\Psi\rangle + (|x\rangle + |\varphi\rangle) = (|\Psi\rangle + |x\rangle) + |\varphi\rangle$ ，

$z(|\Psi\rangle + |x\rangle) = z|\Psi\rangle + z|x\rangle$ ，

1|Ψ〉=|Ψ〉，

0|Ψ〉0，以及z0=0。

[6]　存在一种称为两个矢量的标量积（或内积）的重要运算。它可非常简单地用于表达“单位矢量”、“正交性”和“概率幅”概念。（在通常的矢量代数中，标量积为a　bcosθ，这里a和b为矢量长度，而θ为它们方向之间的夹角。）希尔伯特空间矢量的标量积给出复数。我们把两个态矢量|Ψ〉和|x〉的标量积写作〈Ψ|x〉。存在如下代数规则〈Ψ（|x〉+|φ〉）=〈Ψ|x〉+〈φ|φ〉，〈Ψ（q|x〉）=q〈Ψ|x〉x以及$\langle\Psi|x\rangle=\overline{\langle x|\Psi\rangle}$，在这里横道表明复共轭（z=x+iy的复共轭为$\bar{z}=x-iy$，x和y为实数；注意$|z|^2=z\bar{z}$）。态|Ψ〉和|x〉的正交性表为〈Ψ|x〉=0。态|Ψ〉的长度平方为$|\Psi|^2$=〈Ψ|Ψ〉，这样|Ψ〉归一化成单位矢量的条件为〈Ψ|Ψ〉=1。如果一个“测量的行为”使|Ψ〉跃迁到|x〉或某种和|x〉正交的态，则它跃迁到|x〉的幅度为〈x|Ψ〉，此处已假定|Ψ〉和|x〉都是归一化的。若还没有归一化的话，从|Ψ〉到|x〉的跃迁概率写作〈x|Ψ〉〈Ψ|x〉/〈x|x〉〈Ψ|Ψ〉。（见*Dirac 1947*）

[7]　熟悉量子力学算符形式的读者，这一测量（按狄拉克符号）用有界限的厄米算符|x〉〈x|来定义。本征值1（对于归一化的|x〉）为是，而本征值0表示非。（矢量〈x|，〈Ψ|等属于原先希尔伯特空间的对偶空间。）见*von Neumann 1955*，*Dirac 1947*。

[8]　　在我早先对包含单独粒子的量子系统的描述中，有点过于简略。那时候我不管自旋，而假定只按照它的位置来描述态。实际上存在某些称作标量子的粒子，譬如叫作π子（π介子，参阅281页）的核子或某些原子——其自旋值为零。对于这些粒子（也只有这些粒子）上述只按照位置的描述在实际上是足够的。

[9] 取$|\swarrow\rangle=\bar{z}|\uparrow\rangle-\bar{w}|\downarrow\rangle$，这儿$\bar{z}$和$\bar{w}$是z和w的复共轭。（注释6）

[10]　有一种标准的实验仪器，称作斯特恩——盖拉赫仪的可以用来测量适当的原子的自旋。原子束被射入并通过一个高度非均匀性的磁场，而场的非均匀性的方向为测量自旋提供了方向。原子束被分裂成两束（对于半自旋的原子而言，若是原子具有更高的自旋，则会分裂成多束）。一束给出原子的自旋答案为是，另一束的答案为非。可惜的是，

由于一种和我们目的无关的技术上的原因，使得该仪器不能用于测量电子的自旋。测量电子必须用一种更间接的方法。（见*Mott and Massey 1965*　。）由于种种原因，我宁愿不去特别提及在实际上如何测量电子自旋。

[11] 富有进取心的读者会介意去检验正文中的几何。最容易的办法是把我们的黎曼球面方向调整得使α方向为“向上”，而β方向在由“向上”和“向右”展开的平面上，也就是β方向由在黎曼球面上的q=tan（θ/2）表出，然后用〈x|Ψ〉〈Ψ|x〉/〈x|x〉〈Ψ|Ψ〉来计算从|Ψ〉到|x〉的跃迁概率。参见注释6。

[12] 在数学上我们说，两个粒子的态矢量是第一个粒子的态矢量空间和第二个粒子的态空间的张量积。所以态|x〉|φ〉是态|x〉和态|φ〉的张量积。

[13] 沃尔夫冈·泡利是一位优秀的奥地利物理学家和发展量子力学的杰出人物。1925年，他以假设的形式提出了不相容原理。而对我们现在称作“费米子”的完整的量子力学处理是1926年由极具影响的富有创见的意大利（美国）科学家恩里科·费米和我们已提到过好几次的伟大的保罗·狄拉克发展的。费米子的统计行为按照所谓的“费米——狄拉克统计”，以与可区别粒子的经典统计“玻尔兹曼统计”相分别。玻色子的“玻色——爱因斯坦统计”是由著名的印度物理学家S.N.玻色和阿尔伯特·爱因斯坦于1924年在处理光子时发展的。

[14] 这是一个如此杰出和重要的结果，值得再给出另一种表述。假定在E测量仪中刚好有两个刻度，向上[↑]和向右[→]，而P测量仪中有两个刻度，向右上方45°［↗］和向右下方45°［↘］。E&测量仪和P测量仪实际上分别使用刻度[→]和［↗］来测量。那么两个测量仪相一致的概率为1/2（1+cos135°）=0.146……，比15%稍小一些。用这些刻度进行长系列的试验，譬如得到：

E：是非非是非是是是非是是非非是非非非非是是非……

P：非是是非非非是非是非非是是非是是非是非非是……

“√”“√”“√”

给出刚好低于15%的一致性。我们现在假定P测量不受E刻度的影响——使得如果E的刻度为[↑]而不是[→]的话，P结果也刚好完全一样——并且由于[↑]和［↗］ 之间的角度和[→]和［↗］ 之间的一样，这样在P测量和新的E测量，譬如叫E的测量之间的一致性就又应该刚好比15%低一点。另一方面，如果E刻度和以前一样为[→]，但是P刻度为［↘］ 而不再是［↗］，则

E的结果和以前一样，但是在新的P，譬如称作P'的结果和原先E结果之间的一致性只能刚好比15%低一点。由此推出，如果实际使用这些刻度的话，则在P'测量［↘］ 和E'测量[↑]之间的一致性不会超过45%（等于15%加15%加15%）。但是在［↘］ 和[↑]之间的角度为135°而非45°，因此一致性概率应刚好比85%多一些，而不是45%。这是一个矛盾，它表明E测量的选择不能影响P的结果（或反之）的假定是错误的！我感谢大卫·墨明提供的这一个例子。正文中给出的例子引自于他的文章（*Mermin 1985* ）。

[15] 更早的结果是弗里德曼和克劳塞（1969）在基于克劳塞、霍尼、希莫尼和霍尔特（1969）提出的思想上得到的。在这些实验中还有一点要提到，由于所用的光子探测器的效率比百分之百要低得多，所以在发射出的光子中只有相对少的部分在实际上被观测到。然而，即使用这些效率较低的探测器，测量结果和量子理论的一致性仍是如此完美，很难想象，何以使用更好的检测器会忽然产生比理论更坏的一致性！

[16] 量子场论似乎为不可计算性提供某种新的视界（参见*Komar 1964* ）。

第7章

[1] 一些相对论“纯粹者”宁愿用观察者的光锥，而不用他们的同时空间。然而，这对此结论毫无影响。

[2] 这本书印出后，我发现到那时候两个人都早过世了，只能是他们遥远的后代再回过来“邂逅相遇”。

[3] 在恒星中从轻核子（例如氢核）合并成重核（例如氦核，或最终铁核）的过程会得到熵。同样的，地球上存在的氢中有许多“低熵”，我们总有一天可以利用其中一些，使之在“聚变”核电站中转化成氦。通

过这种手段得到熵的可能性是由于引力已经使得核集中到一起，从而使之离开那些逃逸到浩瀚的空间去的，现在构成2.7开黑体背景辐射的大量光子（参阅410页）。该辐射中包含有比存在于通常恒星中的物质大得多的熵。如果它们完全集中到恒星物质中去，它能用以使大多数这些重核分解为构成它们的粒子！所以在聚变中得到的熵是“暂时的”，引力集中效应的存在才使之成为可能。将来我们会看到，尽管通过核聚变得到的熵和迄今直接通过引力得到的大多数情况相比是非常大——而黑体背景中的熵更巨大得多——这纯粹是局部和暂时的状态。引力的熵源比聚变以及2.7开辐射大到无与伦比的程度（参阅436页）！

[4]　可以把在瑞典的超深的钻井的最近证据解释为对于高尔德理论的支持。但是，该结论是非常令人争议的，还存在另外的传统解释。

[5] 我在这里假定这是所谓的“第Ⅱ类”的超新星。若是“第Ⅰ类”的超新星，我们就再按照从聚变（参阅注释3）提供的“暂时的”熵获得来考虑。然而，类型Ⅰ超新星不太可能产生大量铀235。

[6]　我将具有零或负曲率的模型称为无限模型。然而，存在将这些模型“卷叠”使之成为空间有限的方法。这种不太可能和实际宇宙相关的考虑，不会太影响讨论，我不在此为之忧虑。

[7]　此信念的实验基础主要来自两类数据。第一，粒子以这种相关的速度相互碰撞的行为、反弹、分裂以及产生新粒子。这可从在地球上不同地点建造的高能粒子加速器，以及从由外空打到地球上的宇宙线的行为得知。其次，我们知道制约粒子相互作用方式的参数在10^{10}年内的改变量甚至小于$1/10^6$（参阅*Barrow 1988*）。这样，非常可能的情形是，从太初火球时代开始，它们根本就没有显著地改变过（或可能根本不变）。

[8]　泡利原理实际上不禁止一个电子和另一个电子待在同一“地方”，但是它禁止它们两个处于同一“态”——态牵涉到电子如何运动和自旋。这在实际论证中有一点微妙。它在第一次提出时引起了许多争议，尤其是来自爱丁顿的。

[9]　英国天文学家约翰·米歇尔早在1784年，以及稍后些拉普拉斯亦独立提出这样的论证。他们的结论是，宇宙中大多数的大质量和集中的

物体，正如黑洞那样，也许的确完全看不见。但是他们预言式的论证是利用牛顿理论进行的，因此这些结论充其量只是在某种程度上可使人信服。约翰·罗伯特·奥本海默和哈特兰德·斯尼德（1939）首次提出了适当的广义相对论的处理。

[10] 事实上，在一般的非静态黑洞的情形下，视界的准确位置不是某种可由直接测量确定的东西。它部分地依赖于知道所有物质在其将来都会落入黑洞。

[11] 见*Belinski, Khalatnikov and Lifshitz*（1970）和*Penrose*（1979b）的讨论。

[12] 将引力对系统熵的贡献用整个外尔曲率来测度是诱人的，但何种测度才合适仍不清楚。（一般来说，需要具有某种古怪的非局部性质。）幸运的是，这种引力熵的测度对于现在的讨论不必要。

[13] 现在存在一种众所周知的观点，称之为“暴胀模型”。它的目的是为了解释，譬如讲宇宙在大尺度下的均匀性。根据这个观点，宇宙在其极早期遭受到一个巨大的膨胀——其膨胀数量级比大爆炸模型中的“通常”膨胀大得多。其想法是，任何无规性都被这个膨胀所抹平。但是，如果没有某种更巨大的初始限制，例如由外尔曲率假设所提供的限制，暴胀不会发生。它并没有引入时间不对称的因素，用于解释初始和终结奇性之间的不同。并且它是基于不牢固的物理理论—GUT理论——按照第5章的分类法——其状况并不比尝试类更好些。可参阅彭罗斯（Penrose 1989b）在本章观念的框架中，对“暴胀”的批评。

第8章

[1] 流行的修正包括有：（i）实际上改变爱因斯坦方程里奇=能量（通过“高阶的拉格朗日量”）；（ii）把空间——时间的维数从四维改变成高维（诸如在所谓的“卡鲁查——克莱因类型理论”中）；（iii）引入“超对称”（一种从玻色子和费米子的量子行为中借来的思想），结合成一个更广泛的方案，并将其（不是完全逻辑地）应用于空间——时间坐标；（iv）弦理论（目前流行的激进方案），用“弦历史’来取代“世界线——通常和观念（ii）和（iii）相结合。不管它们是多么流行以及多么有力的表达，所有这些设想，按照第5章的分类，肯定应被归于尝试类中。

[2]　虽然量子化过程不总能保持经典理论的对称性（参阅*Treiman 1985*，*Ashtekar et al.1989* ），这里所需要的是所有四种通常表为T, PT, CT和CPT，对称的破坏。这似乎（尤其是CPT破坏）是在传统量子方法的能力之外。

[3]　就我所知，霍金目前为这事提供量子引力解释的设想中就隐含了这种观点（*Hawking 1987*，*1988* ）。哈特尔和霍金（1983）关于初态量子引力起源的提议也许能给予初始条件外尔=0一些理论实质内容。但是（依我的意见），这些观念目前还没有把根本的时间不对称引入进去。

[4]　按照在第6章注释6给出的标量积的运算〈Ψ|〉来看这些事件将更加透彻。我们在时间向前的描述中计算概率p为

$p=|\Psi|x\rangle|^2=|\langle x|\Psi\rangle|^2$，

而在时间向后的描述中为

$p=|\langle x'|\Psi\rangle|^2=|\langle\Psi'|x'\rangle|^2$。

从〈Ψ'|x'〉=〈Ψ|x〉得出上面两种结果必须是一样的。〈Ψ'|x'〉=〈Ψ|x〉的本质是说“演化”是“幺正的”。

[5]　有些读者也许很难了解，当未来事件已定，要问过去事件的概率能有什么意义！然而，这不是根本的问题。想象描述在空间——时间中的宇宙整个历史。为了求在q发生情况下，p发生的概率，想象考察在所有的q发生的情形下，计算所有这些中伴随有p发生的部分。这就是所需的概率。q是否发生于比p更晚或更早无关紧要。

[6]　这些必须是所谓的纵向引力子——“虚”引力子，它构成了恒定引力场。不幸的是，想以一种清晰明了和“不变的”数学方式来定义这种东西，有些理论性的困难。

[7]　我自己原先计算这个数值的粗略方法被阿拜·阿斯特卡大大改善。我在这里用他的数值（见*Penrose 1987a* ）。然而，他向我强调，在人们似乎必须使用的一些假设中存在大量的任意性，所以在采用所得

到的精确质量值时必须相当小心。

[8] 在文献中时时出现其他不同的尝试，企图为态矢量的缩减提供客观理论。最相关的是卡拉里哈奇（1974），卡拉里哈奇、佛伦克尔和路卡克斯（1986），柯玛（1969），派尔（1985，1988），吉拉地、雷米尼与韦伯（1986）。

[9] 我本人多年来致力于发展称之为“扭量理论”的空间——时间的非定域理论，主要是从其他方向引发的（见*Penrose and Rindler 1986*，*Huggett and Tod 1985*，*Ward and Wells 1990* ）。然而，这个理论至少缺乏某些重要的部分，所以不适合于在此进行讨论。

第9章

[1] 在一次英国广播公司的演讲；见*Hodges 1983* ，419页。

[2] 第一次这类实验是对猫进行的（参考*Myers and Sperry 1953* ）。有关头脑分裂实验的进一步情形，见*Sperry 1966*，*Gazzaniga 1970*，*Mackay 1987* ）。

[3] 关于视皮质功能研究的可读文献，见*Hubel 1988* 。

[4] 见*Hubel 1988* ，221页。更早的实验曾记录到只对一只手的图像有感应的细胞。

[5] 现在已确立的理论认为神经系统由分开的个别细胞（即神经元）组成，它是由伟大的西班牙神经解剖学家拉蒙·卡雅勒在1900年左右大力倡导的。

[6] 事实上，所有的逻辑门都可仅仅由“～”和“&”构成[甚至只要用一个运算～（A&B）就够了]。

[7] 事实上，利用逻辑门比利用第2章内详细考察的图灵机更接近于制造电子电脑。那一章强调图灵方式是基于理论上的原因。杰出的匈牙利/美国数学家约翰·冯·诺依曼和阿伦·图灵的研究对实际电脑的发展的贡献可谓旗鼓相当。

[8]　这些比较在许多方面是误导的。电脑中绝大多数晶体管和“记忆”而不是和逻辑运算有关。而且，电脑记忆总是可以从外界在本质上无限地扩展。随着平行运算的增加，比现在正常情形下更多的晶体管可直接涉及逻辑运算。

[9]　在德义奇的描述中喜欢用量子理论“多世界”观点。但是要紧的是要明白，了解采取什么观点不是主要的。不管人们采用哪一种标准量子力学观点，量子电脑的概念一样合适。

[10]　如果允许“经典”构件为整个齿轮、轴等，则这个评论不适用。我在这里是取通常的（例如，点状或球状的）粒子为构件。

第10章

[1]　我们在第4章153页已经看到，一个形式系统中的证明，检查证明的有效性总是算法的。反过来，任何产生数学真理的算法总能和公理以及平常逻辑（“谓词演算”）的步骤法则相联合，进而为推导数学真理提供一个新的形式系统。数学家当然有时候会犯错误。图灵本人似乎相信，这是哥德尔类型论证反对人类思维为算法的“漏洞”所在。但是依我看来，人类易错不太可能是人类洞察的关键！（而且用算法能非常近似地模拟任意随机者的行为。）

[2]　数学家中的确存在不同的观点，一些读者会被这事实所困扰。回顾一下第4章的讨论。然而，我们不必过于关心它们之间存在的差别。它们只与非常大集合的神秘问题有关，而我们可以把注意力集中于算术命题（具备有限数目的存在和普遍量衡的），这样前面进行的讨论就可以适用。（用于无限集合的反射原理有时可被用于推导算术命题，也许这说得有点过分了一些。）对于甚至宣称不承认存在所谓数学真理而有哥德尔免疫的教条形式主义者根本不予理睬，因为他显然不具备这个讨论中用以判断真理的品质！

[3]　只有到了很晚，大约1968年（多半由于美国物理学家约翰·A.惠勒预言性的思想）“黑洞”这术语才被广泛使用。

[4]　我认为，动物需要睡眠，在睡眠中它们有时显得会做梦（经常可以从狗身上注意到这一点），这个事实是它们具有意识的证据。意识因素似乎是做梦和无梦睡眠之间的差别的重要部分。

[5]　在狭义或广义相对论的情形下，“时间”之处应读成“同时性空间”或“类空表面”（258页，275页）

[6]　然而，在空间无限的宇宙有一个例外，不过在一个人和他紧邻的环境中会有这一切无数的复本（和多世界的情况相当像）！每一复本的未来行为可以稍有不同，而人们将永远不会完全清楚，他实际上是“在”数学模型中的相互近似的哪一个复本中！

[7]　甚至某些实际晶体的成长也会涉及类似问题，例如在所谓的弗兰克——卡斯帕态中，当基本晶胞涉及几百颗原子时。在另一方面，应该指出，小野田、斯特恩哈特、狄温参卓和索科拉（1988）提出了一个五重对称准晶体，在理论上“几乎局部的”（虽然仍然不是局部）的生长步骤。

跋

“……感觉像是？嗯……一个最有趣的问题，我的年轻人。嗯……我自己也想知道答案，”这位总设计师说道，“让我们来看看我们的朋友对这问题怎么说……真奇怪……呃……超子电脑说它不知道……它甚至不能理解你想问的是什么！”会堂里的笑浪声终于爆发成大笑。

亚当感到非常难为情。不管他们做出什么，他们都不应当笑啊！

[1] 在写这类著作时，在没有任何性别含义的地方存在着不可避免地用“他”还是“她”两个代词的问题，在提到某一抽象的人时也遇到了相应的问题。我将用“他”来表明短语“她或他”，这就是我通常所做的。然而，我希望在这儿宁愿用一位女性的质问者这一点“性别主义”能被原谅。我猜想，她或许比她的男性对手对于识别真正的人性会更加敏感些！

[2] 在关于应把什么当成真正地通过图灵检验这一点上，我有点故意奸狡。例如，我可以想象电脑在一长串的失败后，可把人的原先的所有回答放在一起，然后把它们拌上一些随机的成分。过了一阵我们疲倦的质问员也许用完了原先要问的问题，她就可能被愚弄，我把这种方式认为是电脑方面的“欺骗”！

[3] 譬如讲在1989年！

[4] 详情可参阅在第4章结尾处关于复杂性理论和NP问题的讨论。

[5] 事实上，图灵在他原先的描述中允许磁带有更复杂的记号，但这并没有什么本质上的差别。更复杂的记号总能被细分成记号和空白的序列。我将随意地对他原先的详细说明作各种不重要的变通。

[6] 正如这里所描述的，这一流程图本身实际上是“仪器”的一部分。而不是外部环境的“磁带’。我们在磁带上表示的正是实际的数A，B，A-B等。然而，我们还要以一个线性的一维形式来表达该仪器的规约。正如我们将要看到的，和普适图灵机相关的，在一台特殊仪器的规约和对该仪器可能的“资料”（或“程序”）的规约之间有个密切关系。所以，使这两者都处于一维形式是方便的。

[7] 记住我说的自然数是指0，1，2，3，4，5，6，……。我写成“x+1”和“w+3”等，而不写成费马断言的更熟知的形式（$x^w+y^w=z^w$；x, y，z＞0，w＞2）的原因是，我们允许x, w等为从零开始的所有自然数。

[8] 普林斯顿大学的英籍数学家安德鲁·怀尔斯于1993年证明了费马大定理（译者注）。

[9] 我们记得，素数2，3，5，7，11，13，17，……是那些只能分别被它们自己和1整除的自然数。0和1都不认为是素数。

[10] 这是熟知的并有强效的被称为反证法的数学方法。利用这种办法，人们首先假定所要证明的东西是错的.然后从这里推出一个矛盾，就这样证明所需要的结果实际上是对的。

[11] 事实上，由于在上面建造通用图灵机U已使我们能把T_n（n）写成作用于n上的一台图灵机，所以已经得到这个最难的部分。

[12] 一种更熟悉的记号是写成a=b（c），但这些特别的括号不是真正必要的，在习惯上把它们忽略掉更好些。如果把它们一律都保留着，就会导致相当的繁琐，诸如表达式（f（p））（q）

和（（f（p））（q））（r）可分别简化成（fp）q和（（fp）q）r。

[13] 实际上，关于年份的通常惯例并不与此完全相符，这是因为零年被忽略去了。

[14] 注意10^{20}表示100000000000000000000，也就是1后面跟20个0。

[15] 量e=2.7182818285……（自然对数的底。其数学上的重要性可和π相比的无理数）的定义为：e=1+1/1+1/（1×2）+1/（1×2×3）+……，而e^{z}表示e的z次方，e z可展开为：e^{z} =1+z/1+z^{2} /（1×2）+z^{3} /（1×2×3）+……。

[16] 术语“拓扑的”引自一个几何学分支——拓扑学有时称作“橡皮几何学”——在这种几何中，实际的距离是无所谓的，其所关心的只是对象的连续性质。

[17] 部分地根据齐平纳·德尔·费罗和塔尔塔利亚更早的结果。

[18] 正如卓越的阿根廷作家约格·路易斯·波格斯写道的：“……一位著名的诗人更具发现家而非发明家的品格……”

[19] 一个集合表示事物的整体——可被整体地处理的物理对象或数学概念。在数学中，集合中的一个元素（也就是成员）自身经常为一集合，因为集合可以被收集在一起而形成集合。这样人们可以考虑集合的集合以及集合的集合的集合，等等。

[20] 存在以日常术语来表述罗素悖论的十分好笑的方法。想象一个图书馆中有两本目录书，一本目录书刚好列出了所有引用过它们自己的书，另外一本是刚好所有不引用它们自己的书。试问第二本目录书应列到哪一本目录书中？

[21] 虽然费马的全部命题F的真伪性仍然未知，但是个别命题G（0），G（1），G（2），G3S（3），……直到大约G（125000）的真理性是已知的。也就是说，已经知道没有任何一个立方可以是正数立方的和，没有一个四次方为四次方之和等，直到相应的关于125000次方的断言（见77页的译者注脚）。

[22] 当形式系统具有k+1个不同符号加上从未用过的新的“零”时，我们可把字典编序认为是“k+1进位”的自然数的通常顺序。这是因为以零开始的数和这前面的零被略去的同一个数一样。共有9个符号的串的简单字典顺序可以用通常的没有零的十进制写出的自然数得到：1，2，3，4，……8，9，11，12，……19，21，22，……99，111，112，……。

[23] “集合”和“族”之间存在差异，集合可允许集中在一起而形成另外的集合或族，但是族不能允许集中在一起而形成任何种类的更大的聚合。这被认为“太大”了。然而除了这种循环的论述，即集合是那种确能聚集成另外聚合的聚合之外。不存在决定何时聚合可被当作集合或只能被当作族考虑的法则。

[24] 之所以这么称呼直觉主义是因为它反映了人类的思维。

[25] 允许发生这种不幸的可能性实际上是重要的，这样使得能有描述任何算法运算的潜力。我们记得，为了一般地描述图灵机，我们必须允许实际上永远不停止的图灵机。

[26] 事实上，该证明可由一系列步骤组成，这些步骤反映了直到停止以前的机器的动作。机器一旦停止则证明即告完成。

[27] 这是在40页提到的希尔伯特第十问题的负面答案。（例如，参见*Devlin 1988* 。）这里变量的个数是不受限制的。然而人们知道，为了使这种非算法性质成立，实际上需要变量的个数不超过9就可以。

[28] 王浩实际上考虑了稍微不同的问题——用方的花砖，不旋转，并且边缘颜色必须匹配，但是对我们这里这些差别并不重要。

[29] 一个“多项式”实际上是像$7n^4 -3n^5 +6n+15$这样的更一般的表达式，但是这并不增加我们的一般性。当n变大时，任何这类表达式中的所有包含n的更低方次的项都变得不重要（所以在我们的特例中，除了$7n^4$ 项之外可不管其他的项）。

[30] 几乎是这样的，但也不完全；空间探测器行为所需的精度实际上需要在对它们的轨道计算时计入广义相对论效应——存在有能在地球上定位到如此精确（事实上达到几英尺）的仪器，以至于广义相对论的空间——时间曲率效应的确必须考虑在内！

[31] 参阅费恩曼（1985）关于QED理论的通俗解释。

[32] 我在这儿是指大爆炸的“标准模型”，还有许多大爆炸理论的变种，目前最流行的是所谓的“暴胀模型”，依我看来，它无疑是属于尝试的范畴之中！

[33] 尼古拉·伊凡诺维奇·罗巴切夫斯基（1792—1856）是几位独立发明这种和欧几里得几何不同的几何中的一个人。其他人是卡尔·弗里德里希·高斯（1777—1855），菲迪纳德·施维卡德和雅诺斯·波尔约。

[34] 欧多克索斯也是2000年以来行星运动的有用理论的首创者。此理论后来由希帕恰斯和托勒密所发展，之后即被称为托勒密系统。

[35] 用现代的语言，这表明存在分数M/N使得a/b＞M/N＞c/d。只要a/b＞c/d，则在两实数a/b和c/d之间一定存在一个这样的分数，以使欧多克索斯判据确实被满足。

[36] 严格地讲，这仅就将其近似地认为在作匀速直线运动，尤其是没有旋转时而言。地球的旋转的确有（相对小的）可探测到的动力学效应，最明显的即是北半球和南半球的风偏折方式不同，伽利略认为海潮的起因在于这种非均匀性。

[37] 电和磁之间的不同在于单独“磁荷”（亦即北极或南极）似乎不能在自然中分开存在，磁粒子被称作“偶极子”，亦即微小的磁铁（北极和南极连在一起）。

[38] ∂B/∂t在此方程中的存在正是麦克斯韦的理论推导的妙举。本质上讲，方程中的其他所有的项从直接实验证据中都已知道。系数$1/c^2$ 非常小，这正是为何该项未被实验观察到的原因。

[39] 波动方程（或达朗贝尔方程）可写成{$1/c^2$ （∂/∂t）^{2} -（∂/∂x）^{2} -（∂/∂y）^{2} -（∂/∂z）^{2} }φ=0。

[40] 把空间坐标除以c—光速——的原因是为了以后使用便利，使光子世界线和垂直方向的夹角为45°。

[41] 然而，对于由于s^2 负值而分隔开的事件，量c √（s−2）有一种意义，对于看到这两个事件同时发生的观察者而言，它即是通常的距离（参见后面）。

[42] 也许我们可以说，波动方程和麦克斯韦方程类似，（参阅242页的脚注）也是一个相对论性方程。这样，我们早先考虑过的玻——埃勒——里查德“可计算性现象”也是一个只对在S的有界区域中的初始值而言的效应。

[43] 在牛顿理论中，一个粒子的动能为1/2 mv^2 ，此处m为质量，v为速度；但在狭义相对论中，表达式要稍微复杂些。

[44] 该检测不可以干扰粒子通过t点。可将许多探测器放置在围绕着s的其他许多地方，当这些探测器都没有发生咔嗒的声响时，就可推理粒子通过t点！

[45] 由于在一个准确点上找到一个粒子的概率为零，所以在这里产生了技术上的困难。我们把|Ψ（x）$|^2$ 定义为概率密度，它表示在我们定义的点附近的某个很小的固定尺度的间隔内找到该粒子的概率。这样，Ψ（x）定义了幅度密度，而不是一个幅度。

[46] 按照更标准的分析的描述，我们的每一个螺旋（也就是动量态）由表达式$Ψ=e^{px/h}$ =cos（px/h）+isin（px/h）给出（见第3章116页），这里p是问题中动量的值。

[47] 在更通常的量子力学描述中，将此和除以归一化因子——此处为√2 就得到（$Ψ_t$ +$Ψ_b$ ）/√2—但在这里没有必要用这种方式使描述更加复杂。

[48] 大卫·希尔伯特，我们已在前面的章节中提到了他的名字，在量子力学发现以前很久，他在无限维的情况下，并为了完全不同的数学上的目的，引进了这个重要的概念！

[49] 这里必须在允许矢量的无限求和的意义下才行。希尔伯特空间牵涉到了有关这种无限求和的规则（由于这些过于专业性，所以我不详细论及）。

[50] 正如早先的许多不同地方，我宁愿不使用1/√2 之类的因子以免弄乱这些描述。该因子是当我们要求|→〉 和|←〉 归一化时所引起的。

[51]　这个客观性是我们认真采用标准量子力学形式的一个特征。在一种非标准的观点中，系统也许事先已“知道”它将提供给任何测量的结果。还会带给我们物理实在的一种不同的显然客观的图像。

[52] 复数-p和p一样同为q的平方根，并给出同一偏振椭圆。取平方根和光子是自旋为1也就是基本单位$\hbar/2$ 的二倍以及质量为零的事实相关。对于引力子——这种还未探测到的质量为零的量子引力的粒子——自旋为2也就是基本单位的4倍，在上述的描述中我们要取q的四次方根。

[53]　更准确地讲，角动量是由不同数量的点的这种形态的复线性组合所描述。由于在复杂系统中，不同的叠加可得到不同的总自旋值。这只会使总的图像更不4像经典角动量！

[54]　　然而，在两种方程允许的解的类型方面存在一个重大的差别。经典麦克斯韦场必须是实的，而光子态是复的。光子态还必须满足所谓的“正频率”条件。

[55] 1.对于一般的n, m，机会为$C_{10n}^{m}\div C_{n}^{m}=(10n)!(n-m)!/n!(10n-m)!$

[56]　 这里使用的对数是自然对数，亦即对数底为e=2.7182818285……而不是10，但之间的区别是完全不重要的。一个数n的自然对数x=logn是我们为得到n，而将e自乘的指数，亦即e^{x} =n的解（见115页的脚注）。

[57]　当然，绝不是说，我们相空间点将永不再回到更小的区域中去，如果我们等待足够长的时间，它将最终重新进入这些相对细小的体积。（这被称作庞加莱复现。）然而，在大多数情形，时间尺度是不可思议的长。例如，在气体重新进入盒子的1立方厘米的角落，大约需要10 10 26年，这远比宇宙的年龄长得多！我在下面的讨论中将不理睬这种可能性，因为它在实际上和我们讨论的问题无关。

[58]　　 现在关于这个数值仍有激烈的争议，从60亿到150亿年。这些数值比原先埃德温·哈勃在1930年左右最初观察显示宇宙在膨胀之时以为正确的10亿年大了相当多。

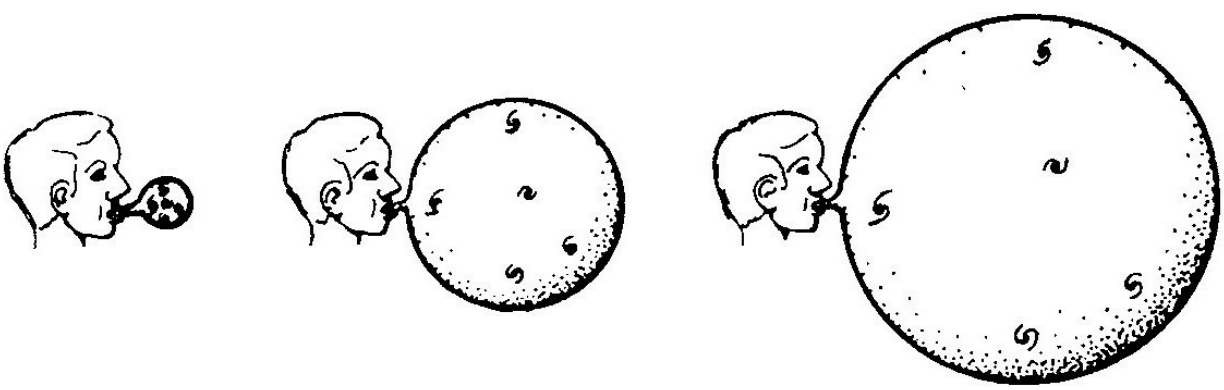


图7.8　宇宙的膨胀可以比喻成被吹胀的气球表面，所有的星系都相互远离

[59] 爱因斯坦于1917年发表了宇宙常数，但在1931年他又撤回，并认为这早年的提议是他的“最

大错误”！

[60] 事实上，在它的最后阶段。白矮星作为一个红色的恒星而发出微弱的光——但是被称作“红矮星”的则是另外具有完全不同特性的恒星！

[61] 我在这些陈述中采用了两个假设。第一是宇宙坍缩比黑洞可能最终消失——这是计入我们将在后面（参阅436页）考虑的、它的（极慢的）霍金辐射引起“蒸发”——更早实现，第二是（非常可能是对的）称为“宇宙监督”的假设（277页）。

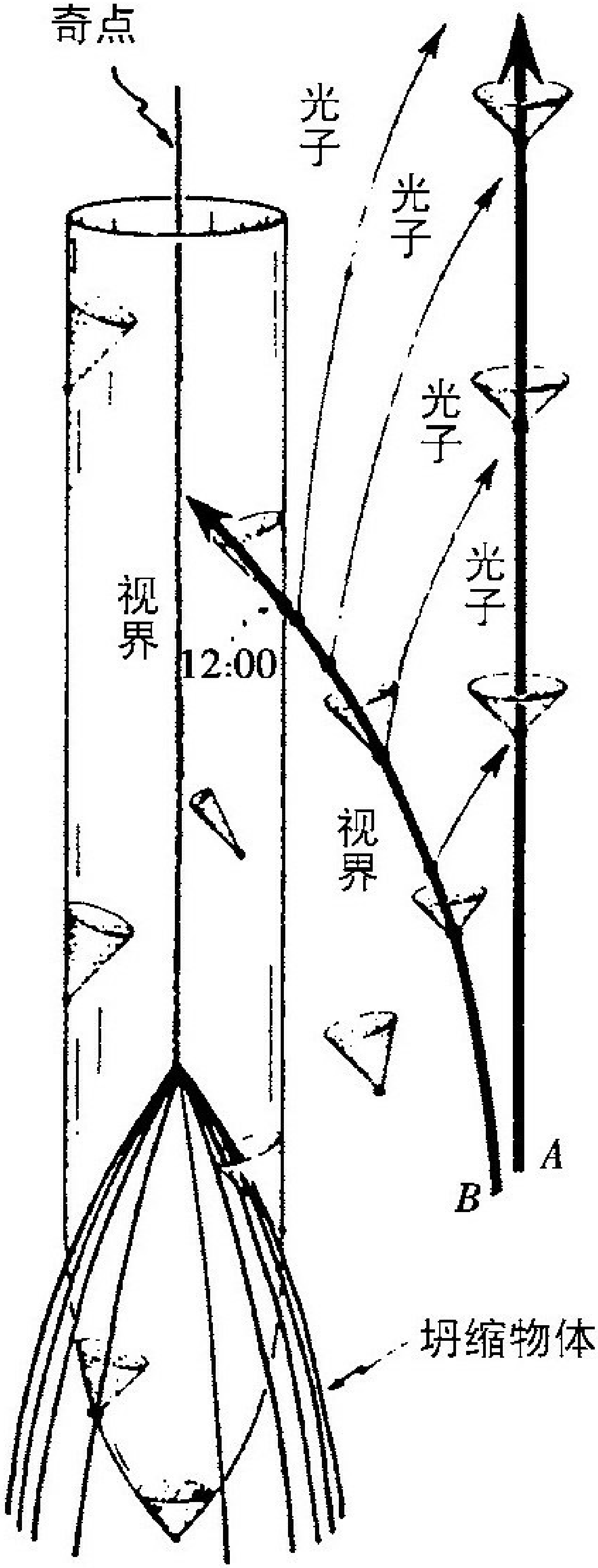

奇点
光子
光子
光子
光子
视界
12:00
视界
A
B
坍缩物体

图7.13　描绘黑洞坍缩的空间——时间图。在图中标作“视界”的为施瓦兹席尔德半径

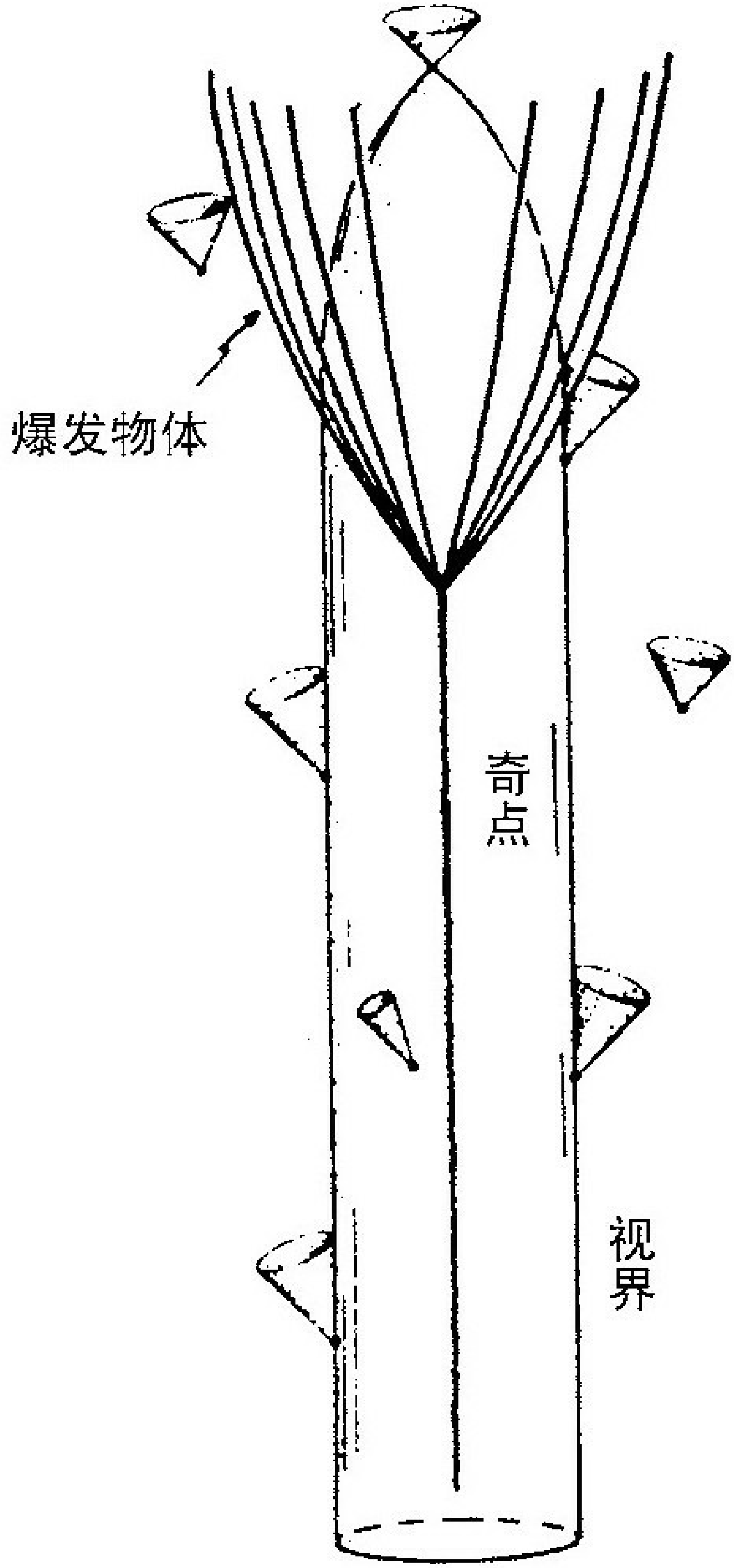

爆发物体
奇点
视界

图7.14　一种假想的空间——时间形态：一个白洞，最终爆发成物质（图7.13的时间反演）

[62] 是这样的一个距离[10^{-35} 米=√（hGc^{-3} ）]，在此尺度下。空间——时间度规本身的“量子起伏”变得这么大，以至于通常光滑的空间——时间连续性的观念不再有效。（量子起伏是海森伯不确定性原理的一个推论——参阅315页。）

[63] 有人会（正确地）争论道，观测的结果不足以清楚支持我主张的宇宙中存在黑洞而不存在白洞的观点。但我的论证基本上是理论性的。黑洞而不是白洞和热力学第二定律相协调！（当然，人们可以直截了当地假设第二定律以及白洞的不存在；但我们试图更深入探究第二定律的起源。）

[64] 奇怪的是，大脑的“交叉行为”不适于小脑。这样，小脑的右半主要是控制身体的右边，而左半控制身体的左边。

[65] 一些令人信服的证据表明，至少黑猩猩能够自我知觉，在允许黑猩猩摆弄镜子的实验显示了这一点，参阅Oakley 1985，第4章和第5章。

[66] 称作“盲视否认”的情况是另一种对盲视的补充。一个事实上全盲的，但坚持他完全能看东西的病人似乎具有推断周围环境的视觉意识！（见*Churchland 1984*， P143。）

[67] 至少对于现代电脑技术而言（参见第1章关于图灵试验的讨论）。

[68] 如果两个算法只是输出一样而实际的计算过程不一样，它们能否被认为是同等的，这又是一个难题。见第2章70页。

[69] 当然“他”是指“她和他”，见第5页的脚注。

[70] 可以指出，至少有一种量子引力论的方法，该方法涉及不可计算性的因素（Geroch and Hartle 1986）。

[71] 在二维空间——时间中这种在时间和空间之间的对称会变得更加显著。二维空间——时间物理的方程式相对空间和时间的交换本质上是对称的。然而，没人会认为在二维物理中空间在“流动”。如果认为在我们知道的物理世界经验中，使时间“实际流动”只是因为在我们空间——时间内恰好空间维数（3）和时间维数（1）之间不对称，那是令人非常难以接受的。

第一推动丛书: 综合系列

The Polytechnique Series

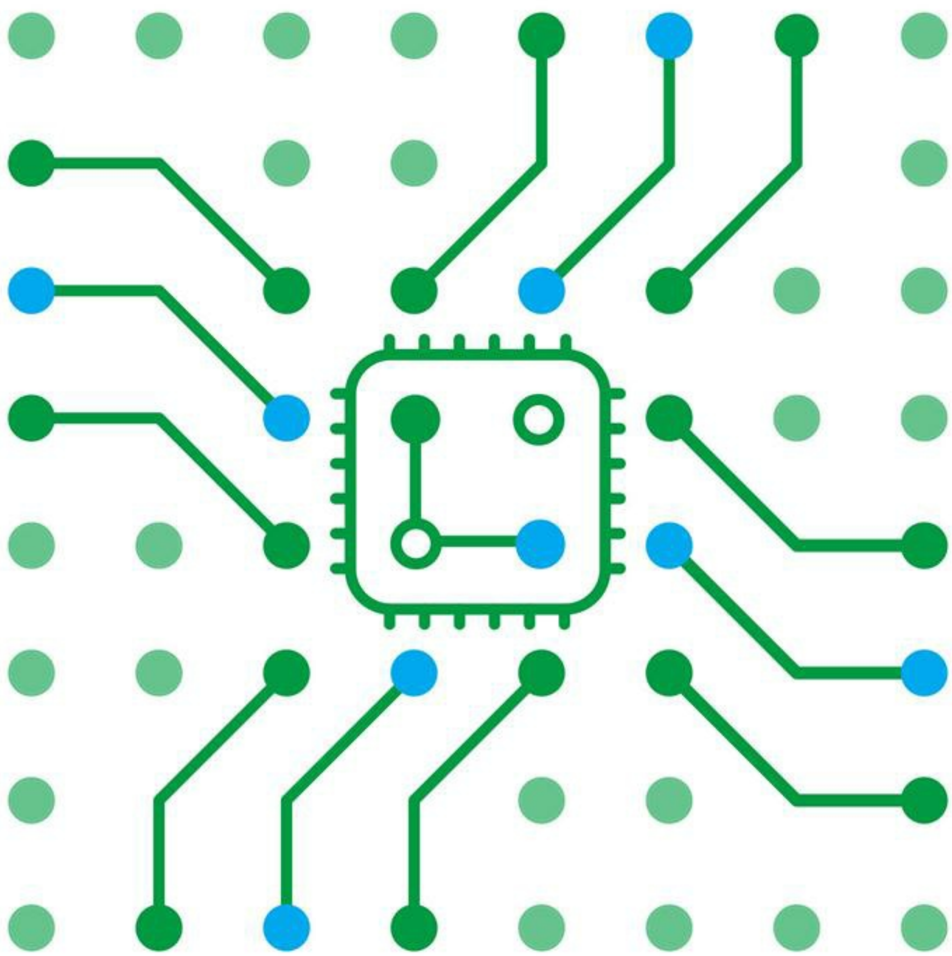
逻辑的引擎

Engines of Logic

[美] 马丁 · 戴维斯 著

张卜天 译

Martin Davis



湖南科学技术出版社

版权信息

逻辑的引擎

著者：[美]马丁·戴维斯

译者：张卜天

责任编辑：吴炜　戴涛　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794420

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
前言
引言
第1章 莱布尼茨之梦
第2章 布尔把逻辑变成代数
第3章 弗雷格：从突破到绝望
第4章 康托尔：在无限中摸索
第5章 希尔伯特的营救
第6章 哥德尔使计划落空
第7章 图灵构想通用计算机
第8章 研制第一批通用计算机
第9章 超越莱布尼茨之梦
尾声
注释
参考书目
译后记

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

前言

本书讲述的是我们的现代计算机背后的那些基本概念和发展出这些概念的人。1951年春，当我在阿兰·图灵（Alan Turing）本人曾于10年前工作过的普林斯顿大学获得了数理逻辑博士学位之后不久，我便在伊利诺伊大学讲授一门以他的思想为基础的课程。有一位一直在听我的讲座的年轻的数学家使我注意到教室的街对面正在建造的两台机器，他认为它们就是图灵观念的物理体现。不久，我就在为这些早期的计算机编写软件了。我持续了半个多世纪的职业生涯便是围绕着现代计算机背后的抽象逻辑概念与它们的物理实现之间的关系而展开的。

计算机从20世纪50年代的塞满整个房间的庞然大物，逐渐演变成今天轻巧而强大的能够完成各种任务的机器，在这整个过程中，其背后的逻辑始终保持如一。这些逻辑概念是几个世纪以来数位天才思想家一步步发展出来的。在本书中，我将讲述这些人的生活故事，并解释他们的部分思想。这些故事本身是引人入胜的，我希望读者们不仅能够喜欢它们，而且在读完之后能够更加了解计算机内部的秘密，同时对抽象思想的价值多一份敬意。

在本书写作过程中，我曾得益于各种各样的帮助。约翰·西蒙·古根海姆纪念基金会在研究的早期阶段提供了热情的经济资助，正是当时所做的那些研究才使本书得以问世。Patricia Blanchette, Michael Friedman, Andrew Hodges, Lothar Kreiser和Benson Mates慷慨地与我分享他们那些专业知识。Tony Sale友好地充当了我游览布莱奇利庄园的导游，图灵曾于二战期间在那里对破译德军的秘密通信起了关键性的作用。Eloise Segal是一位忠实而热心的读者，他帮助我避免了解释方面的缺陷，可惜，他在这本书写成之前就离开了这个世界。我的妻子弗吉尼亚竭力使我避免行文含糊不清。Sherman Stein极为认真地读了原稿，提出了许多改进意见，而且纠正了我的几处错误。我还得益于Egon Börger, William Craig, Michael Richter, Alexis Manaster Ramer, Wilfried Sieg和FransÇois Treves等人的翻译。提出有益建议的其他读者还有：Harold Davis, Nathan Davis, Jack Feldman, Meyer Garber, Dick and Peggy Kuhns和Alberto Policriti。我在W.W.诺顿公司的编辑Ed Barber用他那关于英语散

文的学识慷慨地对本书加以润色，多处改进都直接得益于此。Harold Rabinowitz向我引见了我的代理商Alex Hoyt，后者自始至终都在帮助我。当然，这一长串名字只是要表达我的感激之情，而不是要使我摆脱本书不足之处的责任。有关评论或修正，读者可发邮件至davis@eipye.com，我将心怀感激。

马丁·戴维斯

伯克利，2000年1月2日

平装版注释：当本书的精装版（名为《通用计算机》（*The Universal Computer*））正要付梓刊行之时，Julio Gonzales Cabillon和John McCarthy及时使我注意到了几处潜在的令人难堪的错误。对此我感激不尽，同时也很遗憾他们的帮助未能来得及在前言中提及。就在那本书面世后不久，和我一同游览尼罗河的同伴Connie Holmes使我注意到，有一段曾经被我认为是极为明晰的段落实际上是含糊不清的，我已在这一版中进行了修改。谢谢你，Connie。

引言

如果，一台为了微分方程数值解而设计的机器与百货商店里的一台用来开账单的机器的基本逻辑是一致的，那么我将把这看成我所遇到过的最令人惊异的一致。

——霍华德·艾肯，1956年 [1]

现在，让我们回到理论计算机的类比上来……可以证明，我们能够制造出一台那种特殊类型的机器来完成所有这一切工作。事实上，它可以作为任何其他机器的一个模型。这种特殊的机器或可被称为“通用机”。

——阿兰·图灵，1947年 [2]

1945年秋，正当包含了数千根真空管的巨型计算机ENIAC（电子数字积分计算机）在费城的摩尔电子工程学院接近完成之时，一群专家定期会面以对其计划中的继承者EDVAC（电子离散变量自动计算机）进行讨论。时间一周周过去了，会面时的言辞变得愈发激烈起来，专家们也随之分成了两派，即所谓的“工程师”派和“逻辑学家”派。工程师派的领导者约翰·普莱斯伯·埃克特无可厚非地对他在ENIAC上的成就感到骄傲。要让15000根热真空管共同完成任何一项工作，这都被认为是不可能的。然而，通过小心稳妥的设计原理，埃克特出色地完成了任务。不过，一份关于EDVAC的设计报告正在流传，署名者是小组中逻辑学家派的领导者——著名的数学家约翰·冯·诺依曼（这使埃克特颇为恼火），也正是在此时，争论达到了白热化。那份报告毫不在意工程细节，而是提出了今天以冯·诺依曼结构而闻名的逻辑计算机的基本设计。

尽管ENIAC是工程上的极品，但它却是一堆逻辑上的东西。正是冯·诺依曼作为一个逻辑学家的技能——以及他从英国逻辑学家阿兰·图灵那里学到的东西——使他能够理解，计算机实际上是逻辑机器。它的电路体现了几个世纪以来一大批逻辑学家所提出的观点之精华。当前，正

当计算机技术以惊人的速度前进时，正当我们羡慕工程师们令人瞩目的成就之时，我们很容易忘记那些逻辑学家，正是他们的思想使得这一切成为可能。本书讲述的就是他们的故事。

第1章　莱布尼茨之梦

矿藏丰富的哈尔茨（Harz）山脉位于德国城市汉诺威东南，自公元10世纪起就已经有人来这个地区采矿了。由于地层深处含水较多，所以只有用水泵把水抽到河湾里才能采矿。17世纪时，水车使这些水泵的能力变得强大起来。但不幸的是，这就意味着当冬季水流冻结时，有利可图的采矿工作就不得不终止下来。

1680—1685年，哈尔茨山的矿产管理者开始与一个不易相处的矿工频频发生冲突，这个矿工就是时年30多岁的G.W.莱布尼茨。莱布尼茨是要把风车作为一种额外的能源装置引进来，从而使得采矿工作可以常年进行。此时，莱布尼茨已经取得了许多成就。他不仅在数学上做出了重大发现，而且还以一位法学家而闻名，并且在哲学和神学方面写有大量著述。他甚至还担任了路易十五宫廷中的一项外交职务，以使这位法国的太阳王意识到对埃及（而不是对荷兰和德国）发动一场军事战争的好处。[1]

大约70年前，塞万提斯曾经写了一个忧郁的西班牙人与风车的不幸遭遇。与堂吉诃德不同，莱布尼茨是个顽固的乐天派。面对着世界上显而易见的苦难，莱布尼茨回应那些痛苦万分的人说，上帝对所有可能的世界都无所不知，他无可指责地创造了所有可能世界中最好的一个，我们世界中的一切邪恶因素都以一种最佳的方式为善所平衡。[1] 然而最终的情况表明，莱布尼茨卷入哈尔茨山的采矿项目是极大的失败。他的乐观主义使他没有预见到，内行的采矿工程师会对一个声称要教他们如何做生意的新手抱以天然的敌意，他也没有考虑到风的不可靠性，以及一种新的机器不可避免地需要一个试验阶段。然而最不可思议的乐观想法是，他原本打算能够用他从这个项目中获得的收益开展一些工作。

莱布尼茨的眼光惊人地广阔和宏大。他为微积分运算而发明的符号一直沿用至今，这使得人们不用过多思考就可以很容易地进行复杂的演算。实际进行工作的似乎就是那些符号。在莱布尼茨看来，我们对整个人类知识领域也可实施类似的举措。他梦想对一种普遍的人工数学语言和演算规则进行一种百科全书式的汇编，知识的任何一个方面都可以用

这种数学语言表达出来，而演算规则将揭示这些命题之间所有的逻辑关系。最后，他梦想能够制造出完成这些演算的机器，从而使心灵从创造性的思考中解脱出来。尽管莱布尼茨抱着乐观的态度，但他知道，把这个梦想转变为现实的任务非他个人力量所能及。不过他的确相信，如果有一些有能力的人在一个科学院中共同工作，那么相当一部分任务是可以在若干年内完成的。正是出于为这样一个科学院筹款的目的，莱布尼茨才卷入了哈尔茨山项目。


GODEFROI GUILLAUME
LEIBNITZ,
Né le 3 Juillet 1646 mort le 14 Novembre 1716.

莱布尼茨的奇思妙想

1646年，莱布尼茨出生于德国的莱比锡。那时的德国被分成了1000多个半自治的政治单元，几乎为持续了近30年的战争所毁。30年战争直到1648年才结束，尽管欧洲所有的主要力量都参与了这场战争，但它主要是在德国本土进行的。莱布尼茨的父亲是莱比锡大学的哲学教授，当孩子仅6岁时就去世了。到了8岁的时候，莱布尼茨不顾老师的反对，开始阅读父亲图书馆中的藏书，不久他便能够熟练地阅读拉丁文作品了。

莱布尼茨注定要成为人类历史上最伟大的数学家之一。他从他的老师那里得到了数学思想的启蒙，但老师们对欧洲其他地方的革命性数学著作一无所知。在当时的德国，即便是欧几里得的初等几何也是一门高等学科，人们通常只是在大学阶段才开始学习它。然而当莱布尼茨只有10岁时，他的老师就把亚里士多德于2000年前提出的逻辑系统介绍给了莱布尼茨，这门学科唤起了他的数学才能和激情。莱布尼茨对亚里士多德把概念分成固定的“范畴”着了迷，他产生了一种“奇思妙想”：他想寻求这样一张特殊的字母表，其元素表示的不是声音而是概念。有了这样一个符号系统，我们就可以发展出一种语言，我们仅凭符号演算，就可以确定用这种语言写成的哪些句子为真，以及它们之间存在着什么样的逻辑关系。莱布尼茨一生都沉迷于亚里士多德的理论，并且对此矢志不渝。

事实上，莱布尼茨在莱比锡写的学士论文就是关于亚里士多德形而上学的。他的老师在同一所大学的论文论述的是哲学与法律之间的关系。莱布尼茨显然也被法律研究所吸引，他又获得了一个法律学士学位，这一次他写的论文强调了系统性的逻辑在法律方面的应用。莱布尼茨对数学的第一项真正贡献源于他在大学讲授哲学课程的资格论文（*Habilitationsschrift*）：作为他关于一个概念符号系统的奇思妙想的第一步，莱布尼茨预见到有必要清点这些概念的各种不同组合方式。这使他系统地研究了基本元素复杂排列的数目问题。这方面的工作首先见于他那篇大学授课资格论文，然后是那部内容更加广泛的专著《论组合术》（*Dissertatio de Arte Combinatoria*）。[2]

在继续进行法律研究的过程中，莱布尼茨为获得莱比锡大学的法律

博士学位而提交了一篇论文。它的主题具有典型的莱布尼茨风格，即用推理来解决那些用一般方法难以处理的法律案件。由于种种原因，莱比锡大学并没有接受这篇论文，于是莱布尼茨就把它转交给纽伦堡附近的阿特道夫（Altdorf）大学，在那里这篇论文获得一致好评。22岁那年，莱布尼茨的正式教育完成了，他面临着毕业生的常见问题：如何获得一个职位。

巴黎

莱布尼茨对在德国当大学教授没有多大兴趣，他还有另一条路可走，那就是找一个富有的贵族做资助人。他找到了美茵茨选帝侯的侄子约翰·冯·博伊纳堡，他让莱布尼茨去修订基于罗马民法的法律体系。不久，莱布尼茨被委任为高等上诉法院的法官，同时还参与了一些外交谋略，其中包括未能得逞的对波兰新任国王的选举进行干预，以及前往路易十四的宫廷执行一项任务。

30年战争使得法国成为欧洲大陆的霸主。坐落于莱茵河畔的美茵茨在战争期间就尝到过被军事占领的滋味。因此，美茵茨人非常清楚阻止敌人采取军事行动以及与法国保持良好关系的重要性。正是在这种情况下，博伊纳堡和莱布尼茨才策划说服路易十四及其幕僚意识到把埃及作为军事目标的巨大利益。这一建议——事实上，正是同一建议使拿破仑在一个世纪后陷入了军事灾难——最重大的历史后果就是把莱布尼茨带到了巴黎。

莱布尼茨于1672年来到巴黎，为的是促成埃及计划，并且帮助解决博伊纳堡的一些财政方面的问题。就在这一年，博伊纳堡死于中风的消息传来，这对他是灾难性的打击。尽管莱布尼茨仍在为博伊纳堡家族服务，但却丧失了可靠的收入来源。不过，他设法在巴黎又待了4年，在这4年里他硕果累累，极为多产，其间还对伦敦做了两次短暂访问。[3] 1673年，他在第一次访问时展示了一台能够执行四种算术基本运算的计算机模型，这使他被一致推选为伦敦皇家学会会员。尽管帕斯卡曾经设计过一台能够进行加减运算的机器，但莱布尼茨的机器却可以进行乘除运算，这还是历史上的第一次。[2] 这台机器包括了一个天才的部件——“莱布尼茨轮”，直到20世纪，这一部件仍在计算装置上普遍使用。关于他的机器，莱布尼茨写道：

> 如果要给这台机器以决定性的称赞，那么我们也许可以说，它将使所有那些从事计算工作的人感到喜悦，众所周知，他们就是那些从事金融业务的管理人员、他人财产的管理者、商人、测量员、地理学家、航海家、天文学家……如果只限于科学上的用途，那么古老的几何表和天文学表可以被修正，新表可以被制造出来，利用它们，我们可以测量一切种类的曲线和形体……尽可能地扩充乘法表、平方表、立方表、其他幂次的表、组合表、变分表以及一切种类的级数表是值得的……天文学家们也将不必继续耐着性子进行计算……因为让优秀的人像奴隶一样把大量时间浪费在计算工作上是不值得的，如果使用机器，这些任务就可以被安全地交给任何人去做。[4]

莱布尼茨的机器只能做普通的算术，但他却把握住了机器演算更为深广的含义。1674年，他描述了一种能够解代数方程的机器。一年之后，他为一种机械装置写了相应的逻辑推理，这样就指出了一个目标，即把推理还原为一种演算，并且最终制成能够完成这些演算的机器。[5]

对于时年26岁的莱布尼茨来说，一个至关重要的事件就是他见到了当时居住在巴黎的荷兰大科学家克里斯提安·惠更斯。43岁的惠更斯此前发明了摆钟，并且还发现了土星环。他最重要的贡献——光的波动理论还没有提出。惠更斯认为光是由波构成的，就像石块投入池塘中泛起的波浪传播开来一样。他的想法与伟大的牛顿完全相左，后者认为光是由一串子弹似的微粒流构成的。[3] 惠更斯交给莱布尼茨一份书目，它使这位年轻人很快就了解了当前的数学研究状况。不久莱布尼茨就做出了重要的贡献。

17世纪数学研究的迅速发展主要得益于两项主要进展：

1.处理代数表达式（一般是高中代数的内容）的技巧已经被系统化，这种强大的技巧一直沿用至今。

2.笛卡儿和费马分别论证了，如何通过用一组组的数对表示点而把几何归结为代数。

数学家们都在运用这种新的手段来解决那些以前无法处理的问题。相当一部分工作要涉及极限过程，也就是说，用最终结果的近似值来一

步步地逼近这个结果，从而解决一个问题。其想法是，近似值是不够的，只有逼到极限才能获得一个精确解。

举一个例子便可说明这个概念，它是莱布尼茨早期得到的一个结果，对此他很是自豪：

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\frac{1}{11}+\cdots.$$

等号的左边是我们所熟悉的圆的周长和面积公式中的π，[4] 等号右边则是所谓的无穷级数；它的交替加减的数被称为这个级数的项。省略号的意思是它无限地继续下去。其中的每一项都是以1作为分子，相继的奇数作为分母的分数，它们交替加减，这可以从已经写出的有限的几项看得很清楚：先减去1/11，再加上1/13，再减去1/15等等。但我们真的可以进行无限次加减运算吗？并非如此。但是，从开头一项到任何一项为止，我们都可以获得对“真实”结果的一项近似，随着越来越多的项加入进来，这个近似结果也变得越来越好。事实上，这个近似可以通过加入越来越多的项以达到任意的精确度。对于莱布尼茨的级数来说，这可以从表中显示出来。当我们取到10 000 000项时，得到的数值与π/4的真实值即0.7853981634的前8位保持一致。[5]

项数	精确到小数点后8位的和
10	0.76045990
100	0.78289823
1 000	0.78514816
10 000	0.78537316
100 000	0.78539566
1 000 000	0.78539792
10 000 000	0.78539816

莱布尼茨级数的近似值表

莱布尼茨级数给人留下了深刻的印象，因为它通过一种特别简洁的方式把奇数序列与π这个数继而与圆的面积联系了起来。这是可以利用极限过程来解决的一类问题——确定边界为曲线的图形的面积——的一个例子。另一类可以利用极限来解决的问题是确定准确的变化率，比如一个运动物体不断变化的速度。1675年，莱布尼茨即将结束他在巴黎的逗留，在这一年的最后几个月中，他用极限过程实现了概念和计算上的一连串突破，所有这些工作即被称为他所“发明的微积分”：

1.莱布尼茨发现，计算面积和变化率的问题从某种意义上说很有代表性，因为许多不同种类的问题都可以还原为这两类问题中的某一类。[6]

2.他还认识到，求解这两类问题的数学运算实际上彼此互为逆运算，这在很大程度上就如同加法和减法（或乘法和除法）彼此互为逆运算一样。今天，这些运算分别被称为积分和微分，它们彼此相反这一事实即人们所熟知的“微积分基本定理”。

3.莱布尼茨为这些运算发展出了一套恰当的符号系统（这些符号一直被沿用至今），∫表示积分，d表示微分。[7] 最终，他发现了实际实现

微分和积分所需的数学规则。

这些发现把对极限过程的应用从一种只有少数几位专家能懂的奇特方法，变成了一种可以在教科书中向成千上万的人讲授的直截了当的技巧。[6] 与本书的主题密切相关的是，莱布尼茨的成功使他确信，选取恰当的符号并且定出它们的操作规则是极为重要的。∫和d这些符号并不像一个语音符号系统那样代表着毫无意义的声音，而是代表着概念，这样就为莱布尼茨童年时的那种代表着一切基本概念的符号系统的奇思妙想提供了一个模型。

关于牛顿和莱布尼茨各自完全独立地发明微积分的过程，以及在真相大白之前来往于英吉利海峡两岸之间针对剽窃指控的口诛笔伐，人们已经写过许多著述。对我们的故事而言，莱布尼茨所使用的符号的极大优越性是最重要的。[7] 积分中所使用的一个关键技巧（“置换”法）在莱布尼茨的符号系统中实际上是必然出现的，而在牛顿的符号系统中则更为复杂。甚至有人指出，由于盲目地固守其民族英雄的方法，牛顿的英国追随者们对微积分的发展远远落后于其同时代的大陆同行。

就像许多品尝到了巴黎生活的特殊情趣的人那样，莱布尼茨想在那里尽可能长地待下去。他一方面继续在巴黎工作和生活，一方面又试图维持他在美茵茨的关系。但没过多久，情况就很明朗了，只要他还待在巴黎，美茵茨就不会再给他提供资助。与此同时，他收到了一份发自汉诺威——17世纪德国的许多个公国之一——公爵的职位邀请。尽管约翰·弗里德里希公爵对理智方面的东西不乏某些真正的兴趣，而且还承诺提供经济上的保证，但莱布尼茨并不期望生活在汉诺威。在做出了尽可能长的拖延之后，莱布尼茨的经济状况使他无法继续维持下去，遂于1675年较早的时候接受了这份邀请。在复信中，他要求公爵允许自己能够“在艺术和科学领域中为了人类的利益自由地进行研究”。[8] 1676年秋天，莱布尼茨离开了巴黎，此时巴黎不会再给他提供职位，公爵也不允许更长时间的拖延了。莱布尼茨的余生将一直为汉诺威的公爵服务。

汉诺威

莱布尼茨很清楚，尽管他要求自己能够“在艺术和科学领域中自由地进行研究”，但为了在新职位上取得成功，他不得不去做一些被他的

资助人认为是有用和实际的事情。他着手改良公爵的图书馆，并且提出了各种想法以改善公共管理和农业。在那之后不久，他就开始对哈尔茨山采矿操作进行注定会失败的改进。莱布尼茨所负责的哈尔茨山项目最终被核准了，但仅仅过了1年，公爵于1680年的突然去世使得莱布尼茨的职位受到了威胁。

现在必须要说服新的公爵恩斯特·奥古斯特把莱布尼茨的职位继续下去，并且对哈尔茨山项目进行资助。新的公爵是一个“实际”的人，与其前任不同，他不愿在图书馆上花费过多。莱布尼茨很快就学会了不把恩斯特·奥古斯特卷入学术讨论。为了巩固自己的职位，莱布尼茨提议编写公爵家族的简史。5年之后，当公爵最终叫停哈尔茨山项目时，莱布尼茨提交了一份更为详尽的家族史：如果几处空缺被填补，那么家谱就可以一直追溯到公元600年。公爵显然把这视为雇用历史上一位最伟大思想家的极为合适的方式，他对此毫不吝惜。由于这项任务，莱布尼茨得到了一份固定的薪水、一个私人秘书以及搜集家谱信息的旅费。乐观的莱布尼茨很可能没有想到，自己会在余下的30年里一直为家谱所牵制。（恩斯特·奥古斯特死后，1698年继任的格奥尔格·路德维希对莱布尼茨完成这份家族史尤其不遗余力。）

如果说莱布尼茨在汉诺威有学生的话，那么她们是一些女性，他从未像普通人那样对女性的智力抱有偏见。恩斯特·奥古斯特公爵的很有才华的夫人索菲经常与莱布尼茨就哲学问题进行讨论，而且当莱布尼茨离开汉诺威的时候还与之进行大量的通信。她还设法使自己即将成为普鲁士女王的女儿索菲·夏洛特也能够受益于莱布尼茨的教导。索菲·夏洛特并不单单满足于接受莱布尼茨的智慧，她经常会主动提出一些问题，这有助于莱布尼茨澄清自己的思想。正如当代的莱布尼茨专家班森·梅茨所说：

> 在莱布尼茨一生中的大部分时间里，这些女性是他在汉诺威和柏林的宫廷里的主要支持者。索菲·夏洛特在1705年的突然去世几乎使他崩溃。对他而言，这是一个如此重大的损失，以至于连外国的政府公使都向他表达了正式慰问。当公爵夫人索菲……于1714年去世时，除继续写作不伦瑞克的历史之外，他在其他一切方面的资助都终止了。[9]

这项历史方面的任务确也为莱布尼茨提供了一个外出旅行的借口。

他充分地利用这种自由，这不禁使其资助人大为光火。当然，莱布尼茨还尽可能地与学者保持接触。在柏林，他甚至建立了一个科学协会，后来成为科学院。他的大量通信涵盖了他的兴趣的方方面面。莱布尼茨似乎从不倦于说明，既然上帝在创世方面已经做得足够好了，所以在存在的事物与可能的事物之间必定存在着一种前定和谐，世界上的任何一个事物都有一个充足理由（无论是否能够发现它）。在外交领域，莱布尼茨有两项最受欢迎的计划：重新统一基督教会的各派力量以及为汉诺威公爵获取英国王位的继承权。而当格奥尔格·路德维希果真于1714年（此时距莱布尼茨1716年去世只有2年了）成为英王乔治一世之时，他又粗暴地拒绝了莱布尼茨离开汉诺威的穷乡僻壤去伦敦和自己待在一起的请求，只是命令他抓紧时间完成那份家族史。

普遍文字

那么，莱布尼茨青年时的奇思妙想，即找到一个人类思想的真正的符号系统以及操纵这些符号的恰当的演算工具的宏伟梦想怎样了呢？尽管他不得不承认，如果没有帮助，他就无法在这件事情上取得成功，但他从未忘记这个目标，他终生都在为它进行思考和写作。他很清楚，算术和代数中使用的特殊符号、化学和天文学中使用的符号以及他为微积分运算所引入的符号都提供了范例，说明一个真正合适的符号系统是多么重要。莱布尼茨把这样一个符号系统称为一种文字（characteristic）。与没有实际含义的字母表中的符号不同，在他看来，刚才所说的那些例子都是一种真实的文字，每一个符号都以一种自然而恰当的方式表示某个确定的观念。莱布尼茨认为，我们需要的是一种普遍文字（universal characteristic），即一个不仅真实，而且包含了人类全部思想领域的符号系统。

在一封向数学家G.F.A.洛比达解释这些内容的信中，莱布尼茨写道：代数的“部分秘密就在于文字，也就是说在于恰当地使用符号表达式的技艺”。这种对恰当使用符号的关切就是那根能够引领学者创造出他的文字的“阿里阿德涅之线”[8]。

正如20世纪早期的逻辑学家兼莱布尼茨专家路易·古杜拉所说：

> 可以说，正是代数符号体现了文字的理想，它成了一个典范。

> 莱布尼茨也一直用代数的例子来说明一个恰当选取的符号系统是多么有用，而且是演绎思想所不可或缺的。[10]

也许莱布尼茨对自己所提出的文字的最富激情的说明出现在他致让·伽鲁瓦（莱布尼茨曾与之进行过大量通信）的信中：

> 我对这一普遍科学的用处和实在性越来越深信不疑，我发现极少有人理解它的范围……这种文字是由某种符号或语言构成的……它们完全代表了我们观念之间的关系。这些字符将与迄今为止所想到的字符极为不同。因为人们已经忘记了一条原理，即这种文字的字符将会有助于发明和判断，就像在代数和算术中那样。这种文字将会带来巨大的好处，对我来说，其中有一点尤为重要，那就是用这些字符是写不出在我们看来荒诞不经的想法（chimères）来的。一个无知的人将无法使用它，或者，通过努力学习使用它，他本人将变得博学多才。[11]

在这封信中，莱布尼茨称算术和代数表明了一个恰当的符号系统的重要性。他想到了我们今天仍在使用的以0～9这几个数字为基础的阿拉伯符号系统，对于日常计算来说，它们大大优于早先的系统（比如罗马数字）。当莱布尼茨发现任何数都可以仅仅用0和1表示出来即二进制时，他被这一系统的简洁深深地震撼了。他相信揭示出数的深层性质是有用的。尽管这一信念最终未能证明为合理，但考虑到这种二进制记法与现代计算机之间的关联，莱布尼茨的这种想法是非同寻常的。

莱布尼茨认为他的宏伟计划由三个主要部分组成。首先，在合适的符号被选择出来之前，有必要创造一套涵盖人类知识全部范围的纲要或百科全书。一旦我们完成了这一步，对其背后的关键观念进行选择，并为其中的每一个提供合适的符号就是可能的了。最后，演绎规则可以还原为对这些符号的操作，也就是莱布尼茨所说的“推理演算”（calculus ratiocinator），今天或可称其为一种符号逻辑。在今天的读者看来，莱布尼茨感到无法单凭自己的力量完成这样一个计划是不足为奇的，特别是他还一直受到编写家族史的重压，这被他的资助人视为他的主要任务。然而事情还不只如此，让今天的我们更难理解的是，莱布尼茨如何能够严肃地相信，我们居于其中的纷繁复杂的宇宙可以归结为一种符号演算。

我们只能试着以莱布尼茨看待这个世界的眼光来理解这一点。对莱布尼茨而言，世界上绝对没有任何事物是偶然的或未被决定的；任何事物都遵循着一个计划，上帝对此一清二楚，他正是通过这个计划创造了一切可能世界中最好的世界。因此，世界的一切方面，无论是自然的还是超自然的，都是有关联的，我们可以冀望通过理性的方法来发现这些关联。只有从这种眼光出发，我们才能理解莱布尼茨如何可能在一个著名的段落中说，严肃的“具有善良意志的人们”围坐在桌子旁边来解决某个棘手的问题，在用莱布尼茨所设想的语言——他的普遍文字——写出这个问题之后，人们就可以说：“让我们算一下！”于是人们拿出笔来找到一个解答，其对错必然可以为所有人接受。[12]

莱布尼茨满怀热情地指出了发明“推理演算”这种逻辑代数的重要性，它们对于完成这些演算是不可或缺的：

> 如果对定出正立体数目的人大加褒扬（这没有丝毫的用处，除非是在沉思给人带来愉快的意义上说），如果认为让一个数学天才去揭示一条蚌线、蔓叶线或其他某种几乎没有什么用处的图形是一项值得的训练，那么，让人的推理（这是我们所拥有的最为卓越、最为有用的东西）服从数学定律，这难道不是好得多吗？[13]

虽然莱布尼茨以如此的热情和信心描述了普遍文字，但他却没有完成任何具体的工作。与此不同的是，在创造出一种“推理演算”方面，他的确做出过若干次尝试。他在这方面的部分努力可见于附图。[14] 莱布尼茨提出的逻辑代数是足足超前于他的时代一个半世纪，就像普通代数规定了数字的操作规则那样，这种代数清楚地规定了逻辑概念的操作规则。他引入了一种特殊的新符号来表示把任意多个项组合在一起，这一思想有些类似于把两组事物合成包括两组中所有项的一组事物。他引入了一种特殊的新符号⊕来表示把各项组合在一起。加号启发我们把这种操作看成与普通的加法类似的运算，但它周围的圆圈却警告我们它与普通的加号并不一样，因为相加的并不是数。他的某些代数规则也可见于高中代数课本。从某种程度来说，适用于数的规则也适用于逻辑概念，但还有一些规则是相当不同于那些适用于数的规则的。属于后一类规则的最明显的例子是莱布尼茨的公理2，即A⊕A=A，正是在大体相同的背景之下，乔治·布尔使这条规则成为了其逻辑代数的基础。这条规则说的是这样一个事实，即把若干项与其自身进行组合将不会产生任何新的东西：显然，把某一组事物与同一组事物进行组合，得到的仍将是同

一组事物。当然，数的加法是非常不同的：2+2=4，而不是等于2。

定义3　A在L之内或者L包含A，等价于L可以与以A为其中一项的许多项合在一起后相一致。B⊕N=L表示B在L之内，且B与N共同组成了L。更多数目的项的情形也是一样。

公理1　B⊕N=N⊕B

公设任意多的项，比如A和B，可以被加在一起组成一个单一的项A⊕B。

公理2　A⊕A=A

命题5　如果A在B之内且A=C，则C在B之内。

如果把命题中的A在B之内中的A替换为C，就得到了C在B之内。

命题6　如果C在B之内且A=B，则C在A之内。

如果把命题中的C在B之内中的B替换为A，就得到了C在A之内。

命题7　A在A之内。

因为（根据定义3）A在A⊕A之内，所以（根据命题6）A在A之中。

……

命题20如果A在M之内且B在N之内，则A⊕B在M⊕N之内。

莱布尼茨逻辑演算中的一例

在下一章中，我们将会看到乔治·布尔在对莱布尼茨的努力可能一无所知的情况下，如何沿着莱布尼茨所开辟的方向提出了一种可用的符号逻辑。布尔的逻辑涵盖了亚里士多德2000多年前所引入的逻辑。然而只有到了19世纪，由于戈特洛布·弗雷格的工作，亚里士多德与布尔的逻辑体系所共有的严重的局限性才被真正克服。[15]

尽管莱布尼茨的通信卷帙浩繁，但我们对他个人的情况却并不了解多少。有一位传记作家声称，他在我们所拥有的极少数莱布尼茨的画像中看到了一个疲惫的、不快乐的、悲观的人，这一反他的乐观主义哲学。[16] 其他人则说，他喜欢把糕饼分给邻居的孩子们。据说他50岁的时候曾经向人求过婚，但是当那位女士犹豫的时候他又重新做了考虑，并改变了主意。[17] 我们关于莱布尼茨的印象是：长时间地甚至彻夜坐在书桌旁极为准时地处理大量信件，饭菜则由他的仆人从小饭馆里带给他。他不知疲倦地进行着工作，这一点是毫无疑问的。[9]

如果莱布尼茨没有受到他的资助人的家族史的拖累，并且可以自由地为其“推理演算”花费更多时间，那么情况又将如何呢？他难道不可能完成布尔在很久以后才能完成的工作吗？当然，这种猜测是无用的。莱布尼茨留给我们的是他的梦想，但即使是这个梦想，也使我们对人类的思辨思想充满了敬意，它成为衡量后续发展的一根准绳。

第2章　布尔把逻辑变成代数

乔治·布尔的艰辛岁月

美丽而聪慧的卡洛琳娜·冯·安斯巴赫公主日后将会成为英国王后，即乔治二世之妻。1704年，当她18岁时，她在柏林见到了莱布尼茨。在她随同王室前往英国之后，他们仍然借助通信保持着友谊。她试图说服自己的公公——英王乔治一世——把莱布尼茨带到英国，但正如我们已经知道的，国王坚持让莱布尼茨待在德国完成汉诺威家族史。

卡洛琳娜发现自己被卷入了莱布尼茨与牛顿及其支持者之间的没完没了的愚蠢争论之中，双方都指控对方在微积分的发明上进行了剽窃。她试图使莱布尼茨相信这件事情没有那么重要，但他却不这样认为。事实上，莱布尼茨希望她能劝说国王任命自己做英国的史料编纂者，从而与牛顿担任的造币厂厂长一职相当，而且声称只有这样，与英国相比，德国对抗英国的荣耀才能被保持下来。莱布尼茨给卡洛琳娜写信说，当牛顿声称一颗沙粒能够对遥远的太阳施加一种引力，而无须任何传播这种力的手段时，他实际上是在诉诸神秘的方法来解释一种自然现象，这是无论如何不能接受的。卡洛琳娜则把莱布尼茨的一些著作译成了英文。这一努力使她开始与塞缪尔·克拉克进行接触，有人曾向她举荐后者担当译者。

克拉克是一个哲学家和神学家，也是牛顿的一个忠心耿耿的追随者。在其《上帝的存在与属性》（*Being and Attributes of God*）（1704）一书中，克拉克提出了一种对上帝存在的证明。卡洛琳娜给他看了一封莱布尼茨攻击牛顿观点的信，并要求他做出回复。这使两人之间开始了长时间的通信，直到莱布尼茨去世前几天为止。毫不奇怪，这两人的思想之间没有共通之处。从我们故事的角度看，关于塞缪尔·克拉克的最有趣的事情是，在莱布尼茨去世几乎一个半世纪之后，乔治·布尔将把克拉克对上帝存在的证明作为一个例子，来论证他本人方法的有效性。事实上，通过这些方法，布尔成功地使莱布尼茨的部分梦想焕发了生机，使得克拉克的复杂演绎可以被归结为一组简单的方程。[1]

从莱布尼茨与17世纪的欧洲贵族阶层的世界到乔治·布尔的世界，我们不仅把时间推进了两个世纪，而且还把社会阶层降低了几级。1815年11月2日，乔治出生于英国东部的林肯镇，是四个孩子中的老大。他的父母约翰·布尔和玛丽·布尔在结婚的头9年中一直没有孩子。约翰·布尔是一个补鞋匠，他靠这点生意勉强维持着生活，但却对知识特别是科学仪器有着极大的兴趣。在他的商店橱窗里，他自豪地展出了一架他亲手制作的望远镜。不幸的是，他对生意并不在行，于是支撑整个家庭的重担很快便落在了他那才华横溢的尽职尽责的儿子肩上。[2]

1830年6月，林肯镇的公民目睹了一场无聊的争论。这场争论是在当地的一家报纸上展开的，争论的话题是古希腊作家梅利埃格的一首诗作的英译文的原创性。这篇译文曾作为“林肯14岁的G.B.”的作品刊登在《林肯报》（*LincolnHerald*）上，后来P.W.B.撰文指控G.B.剽窃。P.W.B.承认他无法提供G.B.抄袭的出处，他只是认为这样一篇作品竟会出自一个14岁的孩子之手是不可思议的。这场论战使G.B.和P.W.B.之间来往过几封书信，它们都被原封不动地刊登在《林肯使者》上。

乔治·布尔的家庭早就发现了他的能力，却没有钱让他接受正规的教育。于是，在父亲的重要帮助下，乔治主要依靠自学成才。布尔不仅学习了拉丁语和希腊语，学习了法语和德语，而且还能（当然是很久以后）用这些语言写出数学研究论文。他从未信仰过任何教派，他发现自己不可能信仰基督的神性，但他在整个一生中却秉持着强烈的宗教信念。他不久就抛弃了自己早先成为英国圣公会牧师的念头，这固然是由于他的信仰，但更重要的是因为当父亲的生意破产之后，他的家庭需要直接的经济来源。当乔治在离家40英里以外的一所卫理公会学校当一名教师时，他还不满16岁。2年之后他被解雇了，这显然是因为他的不敬神的行为受到了责难：他星期天研究数学，甚至在做礼拜时也是如此！其实，正是在这个时候，布尔才开始越来越转向数学。后来，他在回忆早年的这段生活时解释说，由于买书的钱非常有限，他发现数学书提供了最好的机会，因为看完它们要比看完其他书花费更长的时间。他还喜欢谈及自己在卫理公会学校期间突然降临到身上的灵感。走过一片田野时，一个想法突然在他头脑中闪过：应该可以用代数形式来表达逻辑关系。这一体验曾被一位传记作家比作保罗走向大马士革的道路，它只有在许多年之后才会结出硕果。[3]



离开卫理公会学校之后，布尔在利物浦找到了一个职位。但在那里教了6个月课之后，他就感到不得不离开了，这可能是因为学校的校长下了逐客令，因为（用他妹妹的话说）他“无所顾忌地沉溺于自己强烈的欲望和激情当中”。[4] 他的下一份工作持续的时间也不长。19岁那年，乔治·布尔决定在他的家乡林肯创办他自己的学校，以使他的家庭得到稳定的经济来源。15年来，布尔一直都成功地担任着校长一职，直到接受了爱尔兰的科克（Cork）城新成立的一所大学的教授职位为止。他的学校（接连有三所）是他的父母和兄弟姐妹唯一的经济支柱，尽管这还需依赖于他的妹妹玛丽·安和弟弟威廉后来的帮助。

虽然经营一所走读的寄宿学校并且讲无数的课程很可能需要整日操劳，但布尔正是在这个时期从一个数学学生转变成了一位富有创造力的数学家。此外，不知怎地，他还抽出一些时间进行社会改良活动。他是林肯镇一个女忏悔者之家的创办人和托管人，其目的是“为在美德的道路上失足的女性提供一个暂时的收容所，通过道德和宗教教育，使她们养成勤勉的习惯，从而赢得社会的尊重”。布尔的传记作家说，这个机构所要帮助的女忏悔者指的就是妓女（维多利亚时期的林肯显然有许多）。[5] 其实更有可能的情况是，这里的顾客通常是一个年轻的女仆，她发现自己怀孕了，同时又被与她处于同一社会阶层的情人在许诺了婚约之后抛弃。：[10] 从乔治·布尔的两篇关于非数学主题的讲演中，我们也许可以对他关于性问题的态度略知一二。在一篇有关教育的讲演中，他警告说：

> 现存的希腊罗马文献中有很大一部分……都被其中提及的异教之罪恶（往往不只是提及）深深地玷污了……但我不相信，当单纯的年轻人面对被恶污染的东西时会没有危险。[6]

在一篇有关适当利用闲暇的讲演中[在“林肯提早打烊协会”（Lincoln Early Closing Association）胜利赢得10小时工作日之后所作]，布尔严厉地说：

> 没有理由在那些背离美德的事务中寻求满足。[7]

同父亲一样，布尔也与林肯的技工学院有着不解之缘。这些技工学

院主要致力于对工匠和其他工人进行业余教育，它们曾在维多利亚时期如雨后春笋般遍及英国全境。布尔在林肯的一家技工学院做一些委员会工作，他的责任是为改善图书馆提供建议、做讲演以及无偿教授许多课程。

然而不知为什么，在做所有这些事情的同时，他还抽出时间研究了英国和大陆的一些最重要的数学文献，并且开始做出自己的贡献。布尔的许多早期工作见证了莱布尼茨对恰当的数学符号系统的力量的信念，符号似乎无须什么帮助就能奇迹般地产生出问题的正确答案，为此莱布尼茨曾举过代数的例子。在英国，当布尔开始自己的工作时，人们已经渐渐认识到代数的力量来自于这样一个事实，即代表着量和运算的符号服从不多的几条基本规则或定律。这就暗示着，同样的力量也可适用于形形色色的对象和运算，只要它们也服从这其中某些同样的定律。[8]

在布尔的早期著作中，他把代数方法应用于那些被数学家称为算子的对象上。它们对普通代数的表达式进行“运算”，以形成新的表达式。布尔对微分算子特别感兴趣，之所以有这样的称呼，是因为它们包含着前一章所提到的微积分的微分运算。[9] 这些算子被认为具有特殊的重要性，因为物理世界中的许多基本定律都具有微分方程（即包含微分算子）的形式。布尔说明了某些微分方程如何可能通过把普通代数方法应用于微分算子而得到解决。今天，工程和科学专业的学生通常在大学二年级或三年级的微分方程课上学习这些方法。

在担任校长期间，布尔在《剑桥数学杂志》（*Cambridge Mathematical Journal*）上发表了不少研究论文。此外，他还提交了一篇很长的论文给《皇家学会哲学会刊》（*Philosophical Transactions of the Royal Society*）。起初，皇家学会不愿考虑这样一个外行所提交的文章，但他们最终还是接受了它，并且授予它金质奖章。[10] 布尔的方法是引入一种技巧，然后把它应用于若干实例。与那些得到解决的例子一样，他通常并不要求证明他的方法是正确的。[11]

就在这个时候，布尔开始与几位顶尖的英国年轻数学家进行通信并且发展了友谊。事实上，正是苏格兰哲学家威廉·汉密尔顿爵士与布尔的朋友奥古斯都·德摩根之间的一场争论，才把布尔的思想带回到了他很久以前的那次灵光闪现，即逻辑关系也许可以表示成一种代数。尽管汉密尔顿在形而上学方面学识很渊博，但他似乎有点像一个热衷于争吵

的愚人。他发表文章对数学作为一门学科进行攻击，这只可能是由于他对这门学科甚为无知造成的。事情的导火索是德摩根发表的一篇关于逻辑学的文章，汉密尔顿声称他剽窃了他本人在逻辑学上的伟大发现，即他所说的“谓词的量化”。我们无须花费时间来理解这一思想或它所引发的激烈争论——它之所以重要，仅仅是因为它激励了乔治·布尔。[12]

曾令年轻的莱布尼茨如此着迷的亚里士多德的古典逻辑包括这样一些句子，如：

1.所有的植物都是有生命的。

2.没有河马是聪明的。

3.有些人说英语。

布尔逐渐认识到，在逻辑推理中，像“有生命的”“河马”或“人”这样一些词的重要之处在于它所描述的所有个体的类（class）或群体（collection）：有生命事物的类、河马的类、人的类。不仅如此，他还认识到这种类型的推理可以用一种关于这些类的代数来表达。布尔用字母来表示类，就像字母以前曾被用来表示数或算子一样。如果字母x和y表示两种特定的类，那么布尔就说，xy表示既在x中又在y中的事物的类。正如布尔本人所说的：

> ……假如一个形容词，比如说“好的”，被用作一个描述词，那么让我们用一个字母，比如说y来表示可以
>
> 用“好的”来描述的所有事物，即“所有好的事物”或“好的事物”的类。再令xy这一组合表示同时适用于x和y所代表的名称或描述词的所有事物的类。于是，如果x表示“白的东西”，y表示“绵羊”，则xy表示“白绵羊”；类似地，如果z表示“有角的东西”……则zxy表示“有角的白绵羊”。[13]

在某种意义上，布尔认为这种类的运算类似于数的乘法运算。然而，他发现了一个重要区别：如果y仍然表示绵羊的类，那么yy表示的是什么呢？它必定表示既是绵羊，又是……绵羊的那种事物的类。但这与绵羊的类是一样的，所以yy=y。如果认为布尔把他的整个逻辑体系都基于如下事实，即当x表示一个类时，方程xx=x总是为真，那么这样说

并不过分。我们以后还会回到这一点上来。：[11]

当乔治·布尔的第一部关于逻辑作为数学的一种形式的革命性专著出版时，他的年龄是32岁。他的阐释更为完善的著作《思维的法则》（*The Laws of Thought* ）出现在7年以后。在布尔的一生中，这段时间是多事之秋。布尔的社会阶层以及不合常规的教育显然使他丧失了在一所英国大学任职的机会。但奇怪的是，正是爱尔兰“问题”给了布尔一个机遇。爱尔兰对英国的许多规定都甚为不满，其中有一项就是他们唯一的一所大学——位于都柏林的三一学院具有新教特色。作为答复，英国政府提议在科克、贝尔法斯特和戈尔韦新建三所大学，称为皇后学院，它们将不受宗教派别限制。尽管受到了爱尔兰政治和宗教知名人士的斥责，因为他们要求建立一所具有绝对的天主教特色的机构，但计划还是向前推进了。布尔决定向这些大学当中的一所申请职位，3年以后，他终于在1849年被任命为科克皇后学院的数学教授。

1849年前后，爱尔兰遭遇了一场由马铃薯晚疫病（一种极具破坏力的真菌疾病，它摧毁了爱尔兰的穷人们赖以生存的马铃薯作物）引起的极为严重的饥荒和病害。许多没有饿死的人也因免疫系统虚弱而被伤寒、痢疾、霍乱和回归热等传染病夺去了生命。英国的统治者很晚才认识到真菌才是这场灾难背后的真正原因，他们反倒指责这是爱尔兰人所谓的好逸恶劳所致。这种社会分析被用来证明从爱尔兰源源不断地出口食物是正当的，而与此同时，数百万人则没有东西吃或被活活饿死。1845年至1852年间，在800万爱尔兰人中，至少有100万人丧命，另外150万人逃往国外。[14]

布尔对此几乎没有什么话可说，他强烈的愤慨集中在对动物的残忍上。事实上，他对爱尔兰人的态度非常暧昧，正如布尔在科克的皇后学院举行落成典礼时所写的一首十四行诗中所说：

你的苦难和眼泪饱经沧桑，

但在智慧上你依然年轻。

你对过去所积存的痛苦思想，

出自那忘却而疲惫的内心。[15]

尽管科克不是主要的知识或文化中心，但这一职位却使布尔过上了一种能够与他作为这个世纪伟大数学家的身份相适应的生活。他的父亲刚刚过世，在为母亲提供适当的必需品之后，他终于可以卸下养家糊口的重担，而考虑过一种个人生活了。对于一所大学来说，在科克的学院里所教授的数学的水平是相当低的。课程提纲以“分数和小数算术”开始，接下来是今天在中学里讲授的那些内容。布尔的年薪是250英镑，每学期还可以从每个学生那里收取2英镑的学费。由于他没有助手，所以他每周要批改他所布置的全部家庭作业。

关于皇后学院的争论仍在继续。虽然科克的校长是著名的天主教科学家罗伯特·凯恩爵士，但天主教显然未被充分地凸显出来：在21个学术成员中，只有他和另一个人是天主教徒。事实上，天主教会的统治集团甚至禁止神职人员参与学院的工作。有些人感到，爱尔兰的职位候选人有时会被有意地忽略，从而为那些相对平庸的英格兰人或苏格兰人创造机会。凯恩校长也不受他的教员们爱戴。由于他的妻子不希望在科克生活，所以这位校长试图从都柏林控制这个学院，加之他的独断专行，这些因素最终酿成了校长与全体教员之间的一场直接斗争，布尔时常被卷入这些毫无结果的斗争当中。[16]

玛丽·埃佛勒斯是布尔未来的妻子，她后来就一些科克居民对她未来丈夫的态度描述了自己的初步印象。当被问及“这位数学教授怎么样”时，一位女士的回答是，“哦，他是那种你可以放心地把女儿托付给他的人”。还有另一位女士，当她向埃佛勒斯小姐解释自己的孩子为什么不在身边时，她说乔治·布尔带他们散步去了，并且说当他和这些孩子在一起时她自己总是很快乐。对于似乎每个人都很喜欢布尔这样一种回答，那位女士提出了异议：

> 他不是我最喜欢的人……至少，我不喜欢与他交往。我不在乎是否与这样的老好人在一起……他从不向你表明他觉得你坏，但是当你接近任何一个如此单纯而圣洁的人时，你不由得会感觉他一定对你感到非常震惊，他使我感到自己非常令人厌恶；但是当孩子们和他在一起时，我总是很自在，我知道他们正在获益。[17]

玛丽·埃佛勒斯是一个性情古怪的牧师的女儿，中校乔治·埃佛勒斯爵士的：[12] 侄女。她也是布尔的朋友和同事约翰·瑞奥的外甥女，后者是科克皇后学院的副校长和希腊语教授，乔治和玛丽就是他介绍认识

的。玛丽从小就表现出了数学上的聪慧。在乔治开始辅导她之后，他们成了好朋友，并且频繁地通信。布尔似乎相信，他们在年龄上相差的17岁可以排除事情进一步发展的任何可能。但在他们第一次会面5年之后，随着玛丽父亲的过世，事情到了该了结的时候了。由于玛丽在经济上难以为继，乔治立即提出求婚。就在这一年，他们结了婚。

他们的婚姻仅仅持续了9年，因为布尔在49岁时就去世了。此前，布尔曾于寒冷的10月在一场暴风雨中步行了3千米去上课。随后患上的支气管炎不久就发展成了肺炎，两个星期之后，他离开了人世。颇具悲剧色彩的是，他妻子对医学的古怪看法可能加速了他的死亡——她似乎用湿冷的床单包裹他来治疗他的肺炎。[18]

显然，这场婚姻是非常幸福的。[19] 玛丽·布尔回忆说，它“就像一场温暖和煦的梦”。布尔的遗孀一直活到20世纪，享年84岁，其时第一次世界大战的战火烧到了海峡对岸。她变得越来越沉迷于各种神秘信仰当中，而且还写了大量不知所云的文字。他们的五个女儿的生活都很有意思。三女儿艾丽西娅的几何能力非常出众，她能够十分清楚地想象四维的几何对象，这使她能够做出一些重要的数学发现。不过，最令人惊讶的还是小女儿埃塞尔·莉莲。当父亲去世时她还只有6岁，她记得自己的童年是在极度贫困中度过的。莉莉（她的昵称）进入了19世纪末生活在伦敦的俄国革命流亡者的圈子里。在一次前往俄国（那时包含了波兰的大部分地方）帮助她那些革命战友的途中，当她凝视华沙的堡垒时，她未来的丈夫威尔弗雷德·伏尼契从牢房中看到了她。数年之后，伏尼契在逃到伦敦时认出了她。这一浪漫的开端促成了他们的婚姻。

莉莉后来以写作《牛虻》（The Gadfly）一书而闻名，这部小说的灵感得自她与一个名叫西德尼·赖利的人之间发生的短暂而热烈的爱情，他不可思议的一生成就了一部名为《赖利：间谍好手》（*Riley: Ace of Spies* ）的电视连续短剧。极具讽刺意味的是，赖利这位强烈反对共产主义的人被布尔什维克在俄国处死，而他的情人的这部真实灵感不为人所知的小说却成了俄国学童的必读书。1955年，《真理报》告诉莫斯科的读者，《牛虻》的作者还活着，在纽约生活，她收到了从苏联寄来的一张15000美元的特许支票。5年后，她以96岁的高龄去世。[20]

乔治·布尔的逻辑代数

回到布尔应用于逻辑的新代数。我们还记得，如果x和y表示两个类，则布尔将用xy表示那些既属于x又属于y的东西的类，他用这个记号是要暗示与普通代数中的乘法的类比。用现在的术语来说，xy被称为x和y的交集。[21] 我们还看到，当x表示一个类时，方程xx=x总是为真。这使布尔提出了一个问题：在x表示一个数的普通代数中，什么时候方程xx=x为真？答案很显然：仅当x为0或1时方程才为真。于是布尔得出了一条原理：如果只限于0和1两个值，那么逻辑代数就成了普通代数。然而，要说明这一点，就必须把符号0和1解释成类。0和1在普通乘法中的运算分别为此提供了线索：0乘以任何数都等于0；1乘以任何数都等于那个数。用符号来表示就是：

$$0x=0，1x=x.$$

对于类而言，如果我们把0解释成一个没有任何东西属于它的类，那么对于任何x，0x都将等于0；用现代的术语来说，0为空集。类似地，如果1包含我们所考虑的每一个对象，那么对于任何x，1x都将等于x，或者说，1是我们所要言说的全体。

普通代数处理的是加减法和乘法。如果布尔要把逻辑代数解释为遵守特殊规则xx=x的普通代数，那么他就需要对+和—做出解释。如果x和y表示两个类，布尔就用x+y来表示或者在x中，或者在y中的所有事物的类，今天它被称为x和y的并集。布尔本人的例子是，如果x表示男人的类，y表示女人的类，则x+y就是由所有男人和女人所组成的类。布尔还用x—y表示在x中但不在y中的事物的类。[22] 如果x表示所有人的类，y表示所有孩子的类，那么x—y就表示所有成年人的类。特别地，1—x将表示不在x中的事物的类，从而

$$x+（1—x）=1.$$

让我们看看布尔的代数是如何工作的。我们用普通代数的记号把xx记作x^2 。于是布尔的基本规则可以写成x^2 =x或x—x^2 =0。根据通常的代数规则把这个方程因式分解，得到

$$x（1-x）=0,$$

用语言来描述就是：没有任何东西可以既属于又不属于一个给定的

类x。对布尔来说，这是一个令人振奋的结果，因为这使他确信自己的路走对了。正如他引用亚里士多德的《形而上学》中的话所说，这个方程精确地表达了

> ……曾被亚里士多德说成是一切哲学的基本公理的“矛盾律”。“同一种性质既属于又不属于同一个东西，这是不可能的……这是一切原理中最确定无疑的……因此，那些做论证的人把这当成一条最终的意见。因为它就其本性而言是其他一切公理的来源。”[23]

当布尔引入新的一般观念时，他一定像所有科学家那样很高兴看到它能够获得证实：像亚里士多德的矛盾律这样一个早期的重要里程碑原来只不过是新观念的一个特殊应用而已。事实上，在布尔的时代，研究逻辑的人普遍都把整个学科等同于亚里士多德在许多个世纪以前所做的工作。正如布尔所说，这就等于坚持“逻辑科学不会像所有其他科学那样有不完美之处，也不会像它们那样有所进步”。亚里士多德所研究的那部分逻辑处理的是一种被称为三段论的非常特殊的受限推理。这些推理从一对被称为前提的命题出发，得出另一个被称为结论的命题。前提和结论都必须用下列四种类型之一的句子来表示：

句子类型	例子
所有X都是Y。	所有的马都是动物。
没有X是Y。	没有树是动物。
有些X是Y。	有些马是纯种马。
有些X不是Y。	有些马不是纯种马。

下面是一个有效的三段论的例子：

所有X都是Y

所有Y都是Z

所有X都是Z

说这个三段论是有效的，就意味着无论X、Y、Z被代之以什么样的性质，只要两个前提为真，那么结论也将为真。下面是这个三段论的两个例子：

所有的马都是哺乳动物。	所有怪物都是蛇鲨。
所有哺乳动物都是脊椎动物。	所有蛇鲨都是紫色的。
所有的马都是脊椎动物。	所有怪物都是紫色的。

我们可以很容易地用布尔的代数方法来证明这个三段论是有效的。说X中的每一个东西也属于Y，就等于说没有任何东西是属于X却不属于Y的，也就是说X（1-Y）=0，或者X=XY。类似地，第二个前提可以被写成Y=YZ。利用这些方程我们得到

X=X Y=X（Y Z）=（X Y）Z=X Z，

即为我们所想要得到的结论。[24]

当然，并不是每一个三段论都是有效的。只要把前面这个例子的第二个前提和它的结论互换一下，我们就得到了一个无效三段论的例子：

所有*X*都是*Y*

所有*X*都是*Z*

所有*Y*都是*Z*

这时就不能用X=XY和X=YZ的前提来得到Y=YZ的结论了。

回想起来，我们很难理解人们竟然普遍相信三段论推理便构成了逻辑的全部。布尔严厉批评了这一观念。他指出，许多日常推理都涉及他所谓的二级命题，即表达其他命题之间关系的命题。这种推理就不是三段论。

举一个这种推理的简单例子。让我们听一下乔和苏珊之间的一场对话。乔找不到他的支票簿了，苏珊正在帮他。

苏珊：你是不是在买东西时把它忘在超市了？

乔：没有，我给他们打过电话，他们没有找到。如果我把它忘在那了，他们肯定就能找到。

苏珊：等一等！你昨天晚上在饭馆开过一张支票，我看见你把支票簿放在夹克口袋里了，如果你从那时起就没有再用过它，那么它肯定还在那里。

乔：你说对了。我没有用过它。它就在我的夹克口袋里。

乔看了看，（如果逻辑学对这一天来还适用的话）丢失的支票簿就在那里。让我们看看布尔的代数如何能被用于分析乔和苏珊的推理。

在他们的推理中，乔和苏珊处理的是如下命题（每一个命题都用一个字母来表示）：

L=乔把支票簿忘在了超市。

F=乔的支票簿在超市找到了。

W=乔昨晚在饭馆开了一张支票。

P=昨晚开了支票之后，乔把支票簿放在了他的夹克口袋里。

H=乔从昨晚起就没有再用过他的支票簿。

S=乔的支票簿仍然在他的夹克口袋里。

他们用了如下推理形式：

前提：

如果L，那么F。

非F。

W且P。

如果W且P且H，那么S。

H.

结论：

非L。

S.

就像亚里士多德的三段论一样，这一形式构成了一个有效推理。因为一如任何有效的推理，被称为结论的句子的真可以从其他被称为前提的句子的真推导出来。

布尔发现，适用于类的同一种代数也可适用于这种推理。[25] 他用一个方程，比如说X=1，来表示命题X为真。于是，他会用方程X=0来表示X为假。因此，对于“非X”，他可写为X=0，而用方程XY=1表示“X且Y”。之所以如此，是因为恰恰当X和Y均为真时，X且Y才为真；而在代数上，如果X=Y=1，则XY=1，但如果X=0或Y=0（或两者都等于0），那么XY=0。

最后，“如果X，那么Y”这一陈述可以用以下方程来表示：

X（1-Y）=0.

为了理解这一点，把这条陈述当成是在断言

如果**X=1**，那么**Y=1**。

而事实上，把X=1代入这一方程，便可得到1-Y=0，即Y=1。

利用这些思想，乔和苏珊的前提就可以用以下方程来表示：

L（1-F）=0，

F=0，

WP=1，

WPH（1-S）=0，

H=1.

把第二个方程代入第一个，我们即可得到L=0，即所要得到的第一个结论。把第三个方程和第五个方程代入第四个，我们就得到1-S=0，即所要得到的另一个结论S=1。

当然，乔和苏珊不需要这种代数。但在人类日常交流背后正在不自觉地发生的那种推理，却可以为布尔的代数所把握。于是鼓励人们期望更为复杂的推理也可能被把握。也许可以认为数学系统地概括了极为复杂的逻辑推理，所以要想对一种以完备性为目标的逻辑理论进行最终的检验，就要看它是否包含了一切数学推理。我们将在下一章回到这个问题。

作为布尔方法的最后一个例子，我们回到这一章开头提到的塞缪尔·克拉克对上帝存在性的证明。我们姑且不论克拉克冗长而复杂的演绎，至少看看布尔是怎么做的就很有意思。我们引用一个小片段：[26]

前提是：

第一，某种东西存在。

第二，如果某种东西存在，那么或者某种东西一直存在，或者现存的东西是从无中产生的。

第三，如果某种东西存在，那么它或者是依其自身本性的必然性而存在，或者是凭借另一个存在者的意志而存在。

第四，如果它是依其自身本性的必然性而存在的，那么就有某种东西是一直存在的。

第五，如果它是凭借另一个存在者的意志而存在的，那么现存的东西是从无中产生的假说就为假。

我们现在必须用符号表达上述命题。

设

x=某种东西存在，

y=某种东西一直存在，

z=现存的东西是从无中产生的，

p=它依其自身本性的必然性而存在（即上面讲到的某种东西），

q=它通过另一个存在者的意志而存在。

布尔接着由前提获得了以下方程：

1-x=0，

x{yz+（1-y）（1-z）}=0，

x{pq+（1-p）（1-q）}=0，

p（1-y）=0，

qz=0.

对于这种把复杂的形而上学推理还原为简单方程的运算，我们不知克拉克会做何感想。作为牛顿的一个追随者，他也许会高兴。但憎恶数学的好斗的形而上学家威廉·汉密尔顿爵士必定会对此惶恐万分。

布尔与莱布尼茨之梦

布尔的逻辑体系不仅包含了亚里士多德的逻辑，而且还远远超过了它。但这距离实现莱布尼茨的梦仍旧非常遥远。考虑下面这个句子：所有失败的学生或是糊涂的或是懒惰的。有人也许会认为这个句子属于

所有**X**都是**Y**

的类型。但这就要求糊涂学生或懒惰学生的类被当成一个单元来处理，而不允许有推理对那些由于糊涂而失败的学生和由于懒惰而失败的学生进行区分。在下一章中，我们将会看到戈特洛布·弗雷格的逻辑体系是如何包含这种更为微妙的推理的。

把布尔的代数用作一个演算规则的系统是非常直接的，我们也许可以说，在其界限之内，它提供了莱布尼茨曾经寻求的微积分的推理演算。莱布尼茨关于这些内容的著述是以书信的形式保留下来的，此外还有其他一些未发表的文献。只是到了19世纪末，才有人认真地搜集和出版这些著作，所以布尔是不大可能知道他的前辈的努力的。不过，把布尔成熟的系统与莱布尼茨零碎的尝试比较一下是有趣的。我们在第1章

中所引述的莱布尼茨的话把A⊕A=A作为第二条公理，于是，莱布尼茨所考虑的运算就会遵循布尔的基本规则：xx=x。此外，莱布尼茨提议把他的逻辑表示成一个成熟的演绎系统，在其中所有规则都可由少数公式推导出来。这与现代的做法相一致，它显示在这个方面莱布尼茨曾超前于布尔。

乔治·布尔的伟大成就是一劳永逸地证明了逻辑演绎可以成为数学的一个分支。尽管在亚里士多德的先驱工作之后，逻辑学上曾经有过某些发展（特别是希腊化时期的斯多葛派和12世纪的欧洲经院学者），但布尔却发现这门学科从本质上说仍然是2000年前亚里士多德之后的样子。从布尔开始，数理逻辑就一直处于连续不断的发展之中。[13]

第3章　弗雷格：从突破到绝望

耶拿是一座具有中世纪风格的小城，后来成为德意志民主共和国的一部分。1902年6月，年轻的英国哲学家伯特兰·罗素往这里发了一封信，收信人是53岁的戈特洛布·弗雷格。尽管弗雷格相信自己已经做出了重大的基本发现，但他的工作却几乎完全被忽视了。当他读到下面这些话时，他一定是比较欣慰的，“我发现我在一切本质方面都赞成您的观点……我在您的著作中找到了在其他逻辑学家的著作中不曾有过的探讨、区分和定义”。但这封信又接着写道：“我只在一个地方碰到了困难。”弗雷格不久就认识到，正是这个“困难”几乎导致他毕生的工作毁于一旦。罗素又继续写道：“对基本问题进行严格的逻辑处理依然相当滞后，我在您的著作中找到了我所知道的这个时代最优秀的东西，因此，请允许我向您表达深深的敬意。”但这已经于事无补了。

弗雷格马上给罗素回信承认这个问题。当时，他把自己的逻辑方法应用于算术基础的著作的第二卷马上就要出版，于是他连忙加了一个补遗，开头是这样的：“正当工作就要完成之时发现那大厦的基础已经动摇，对于一个科学工作者来说，没有什么能比这更为不幸的了。伯特兰·罗素的一封信使我置身于这样的境地。”

又过了许多年，此时距弗雷格去世已逾40年，伯特兰·罗素写道：

> 每当我想到正直而又充满魅力的行动时，我意识到没有什么能与弗雷格对真理的献身相媲美。他毕生的工作即将大功告成，其大部分著作曾被能力远不如他的人所忽视。他的第二卷著作正准备出版，一发现自己的基本设定出了错，他马上报以理智上的愉悦，而竭力压制个人的失望之情。这几乎是超乎寻常的，对于一个致力于创造性的工作和知识，而不是力图支配别人和出名的人来说，这有力地说明了这样的人所能达到的境界。[1]

当代哲学家迈克尔·达米特的大部分工作都受到了弗雷格思想的启发。然而，当他谈到弗雷格的“正直”时，却是别有一番心情：

> 对我来说不无讽刺意味的是，这个曾让我花费大量时间研究其思想的哲学家，至少到了晚年，却是一个恶毒的种族主义者，特别是一个反犹主义者……[他的]日记显示，弗雷格曾经是一个极端右翼分子，他强烈抵制议会体制、民主主义者、自由主义者、天主教徒、法国人和犹太人，他认为他们应当被剥夺政治权利，最好是被逐出德国。我被深深地震撼了，因为我曾经把弗雷格尊为一个绝对理性的人。[2]

弗雷格的贡献极为重要。他提出了把普通数学中一切演绎推理都包含在内的第一个完备的逻辑体系，他用逻辑分析工具来研究语言的开拓性工作为哲学的主要发展提供了基础。今天，在一个标准的大学图书馆中，在“戈特洛布·弗雷格”的条目下可以查到50多项。然而，当他1925年去世时却很痛苦，他认为自己一生的工作毫无结果，他的死也没有受到学术界的注意。即使是现在，我们对他个人生活的了解也少得可怜。[3]

1848年11月8日，戈特洛布·弗雷格出生在注定要成为德意志民主共和国一部分的小城维斯玛。他的父亲是福音派新教神学家，担任一所女子中学的校长职务（他的母亲也在那里工作）。弗雷格38岁那年与35岁的玛格丽特·丽瑟贝格结了婚，7年后丽瑟贝格便去世了，没有留下子嗣。1908年，弗雷格应一位牧师亲戚的请求，领养了一个5岁大的孤儿。正是这个儿子阿尔弗雷德把弗雷格写于1924年即去世前一年的声名狼藉的日记公之于众，从而使迈克尔·达米特对事实的真相气愤不已。阿尔弗雷德·弗雷格本人则于1944年6月在德军占领巴黎的一次战斗中被杀，此时联军登陆诺曼底刚刚过去一个多星期，距巴黎解放也只有两个月了。阿尔弗雷德从他父亲的手稿中打印出这本日记，并于希特勒掌权5年后的1938年把它送到了海因利希·肖尔茨负责的弗雷格档案馆。在那个年代，令迈克尔·达米特气愤不已的情绪在德国似乎是司空见惯的。手稿本身以及阿尔弗雷德写的关于他父亲的一本传记都遗失了。

弗雷格进入大学时是21岁。在耶拿大学待了2年之后，他到了哥廷根大学。3年之后，他在那里拿到了一个数学博士学位。接着，他被任命为耶拿大学的编外讲师。这个时候，弗雷格似乎是由他的母亲资助的，父亲死后，母亲便接管了女子学校。5年以后，弗雷格被任命为耶拿大学的副教授，在那里他一直待到1918年退休。他的同事并不很欣赏他的工作，所以他从未晋升为正教授。

1873年，新统一的德国一片欢腾。与拿破仑三世的法国进行的战争以伟大的胜利而告终，工业正以极快的速度发展着。德皇威廉一世去世之前，他的首相俾斯麦一直通过一种圆滑的政策，即一个精心培育的联盟体系来维护德国的安全。在弗雷格的一生中，俾斯麦和威廉皇帝一直是他心目中的英雄。然而，俾斯麦是一个主张皇帝完全掌管军事外交的彻底的保守分子，他憎恶民主，并且制定法律禁止了社会民主党的许多活动。

威廉二世在继任后不久便解除了俾斯麦的职务。新皇帝是一个虚荣而缺乏自信的人，他实行了一种灾难性的外交政策。他对自己的策略所能造成的影响一次次地判断失误，这使得其他欧洲政权如此恐慌，以至于法国、俄国和英国组成了一个联盟来对付德国。面对着东边俄国和西边法国的腹背受敌的战争危险，德国参谋部抛出了机智但最终是灾难性的“施利芬计划”，以期能够抢在俄国动员起来之前迅速击溃法国。[4]

1914年夏，由于斐迪南大公被刺以及德国的怂恿，奥地利入侵塞尔维亚，挑起了第一次世界大战。为了表明不能允许奥地利侵略其斯拉夫兄弟的决心，俄国开始了动员。德国将军向德皇解释说，作为回应，他们必须马上实施通过比利时发动袭击的“施利芬计划”。随后，德军对中立的比利时的侵犯把英国也拖入了这场灾难性的战争，其后果所造成的阴影笼罩了整个20世纪。战争期间，事情很少按照计划进行，当“施利芬计划”逐渐被终止时，战争陷入了悲惨的僵局，整整一代欧洲人的中坚力量在战壕中惨遭杀戮。德国的学术界中有许多人似乎还不知道战争的形势越来越糟，他们还在为和平奔走呼号，而德国也将并吞大量地盘，包括比利时全境。



正当胜利与德国无缘，英国的包围造成德军伤亡之时，军事指挥权落到了鲁登道夫将军手里。这个反复无常的赌徒（他后来参加了希特勒的啤酒馆政变）拒绝妥协，直到英军在巴尔干半岛的一次突围差点卷击德军的侧翼，鲁登道夫才愁眉苦脸地向德皇禀报不得不休战。这样便结束了战争和德国的君主制。

新的德国共和国的执政党是社会民主党，许多德国人（包括弗雷格）都相信德国是违背自己的意志被迫进行战争的，它不曾战败，而是被社会主义者和（许多人马上又加上）犹太人出卖了。正是这种有害的气氛最终为希特勒掌权创造了机会。

1923年，战后的恶性通货膨胀使私人存款以及弗雷格的养老金变得一文不值。由此导致的贫困迫使他不得不寄宿在亲戚家里，一直到他1925年在维斯玛附近的巴特克莱纳去世。正是在这样的环境下，他写了他那可悲的日记。他希望能有一位伟大的领袖把德国从它所陷入的卑下地位中拯救出来。弗雷格曾经对鲁登道夫抱有很高的希望，但后来却失望地发现他参与了希特勒的政变。他仍然对鲁登道夫将军抱有幻想，但却担心他的年纪太大了。弗雷格没有活着看到兴登堡把共和国的钥匙交给了阿道夫·希特勒。

在1924年4月所写的日记中，弗雷格回忆起一段他家乡的犹太人正以一种他认为合适的方式被对待的时光，而且还公开了他关于法国人及其有害影响的观点：

> 那时有法律规定，犹太人只有在某些一年一度的交易会上才能在维斯玛过夜……我猜想这一条款已经实施很久了。想必是维斯玛的老居民与犹太人接触的经历，才导致他们立了这个法。
>
> 这必定起因于犹太人做生意的方式，以及与这种做生意的方式紧密相关的犹太人的民族特征……甚至连犹太人也有了普选权，甚至连犹太人也有了迁徙自由，这都是来自法国的礼物。我们使法国人用礼物祝福我们变得太容易了。只要想想那些高贵而爱国的德国人……事实上，1813年以前，法国人把我们欺负得够可以的了，而我们却盲目崇拜法国的一切事物……只是在近些年来，我才真正理解了反犹主义。如果一个人希望针对犹太人立法，那么他就必须能

够指定一种区分的标记，从而使人能够确切无疑地识别出犹太人。我一直把这看作一个问题。[5]

充分准确地定义犹太人，以便能够制定法律对付他们，这个问题对于弗雷格来说还只是个理论问题，而在纳粹统治之下就成了一个非常实际的问题。根据这种纳粹的种族规则，20世纪最伟大的思想家之一，并且是弗雷格的崇拜者和学生的路德维希·维特根斯坦就算得上是一个犹太人。

还有些日记责骂了社会民主党人和天主教徒：

> 1914年，帝国遭受着一个毒瘤的侵袭，那就是社会民主党。
>
> （4月24日）
>
> 诚然，我认为教皇绝对权力主义及其在中央党中的体现对我们的帝国和民族是非常有害的；然而，鲁登道夫在他[最近的]文章中对教皇绝对权力主义者的阴谋的揭示深深地困扰着我。[14] 我请求任何还不相信中央党的彻底非德意志精神的人阅读和反思一下鲁登道夫阁下的这篇文章……这是暗中破坏俾斯麦帝国的最邪恶的敌人……[教皇绝对权力主义者]将总是依凭教皇来获得指示。（4月26日）[6]

弗雷格极端右翼的思想在一战后的德国并不罕见。然而，我们也许想知道日记是否只代表着一个痛苦的（或许也是衰老的）老人在临死前一年的思想。可惜，弗雷格曾经持右翼思想一段时间，这是没有什么疑问的。弗雷格的同事，耶拿大学的哲学教授布鲁诺·鲍赫曾在战争期间创立了一个右翼的哲学协会（DPG），并担任其会刊的编辑工作。弗雷格是DPG的早期追随者之一，而且还在其会刊上发表文章。鲍赫关于民族概念的文章宣称，没有犹太人可能真正成为德国人。当纳粹1933年上台时，他的组织曾全力支持过纳粹。[7]

弗雷格的概念文字

当我们从弗雷格晚年可怕的观点转向他年轻时做出的辉煌贡献时，

不禁会大舒一口气。1879年，[15] 他出版了一本不到100页的小册子，名为《概念文字》（*Begrifsschrift* ）。这是一个很难翻译的词，由弗雷格从德文词“Begriff”（概念）和“Schrift”（大致为“文字”或“书写方式”之意）拼接而成。其副标题是：“一种模仿算术语言构造的纯思维的形式语言。”这部著作被誉为“也许是自古以来最重要的一部逻辑学著作”。[8]

弗雷格试图找到一个能够包含数学实践中全部演绎推理的逻辑系统。布尔把普通代数作为出发点，用代数符号来表示逻辑关系。就像其他的数学分支那样，由于弗雷格想以他的逻辑为基础而把代数构造出来，所以引入自己的特殊符号来表示逻辑关系以避免混乱是很重要的。布尔曾经认为用于表示其他命题之间关系的命题是二阶命题，在这里，弗雷格发现那些连接命题的关系也可被用于分析命题的结构，他把这些关系充当了他的逻辑的基础。后来这一重要思想被普遍接受，成为了现代逻辑的基础。

例如，弗雷格会用逻辑关系如果……，那么……来分析句子

所有的马都是哺乳动物。

即

如果x是一匹马，那么x是一个哺乳动物。

类似地，他会用逻辑关系……且……来分析句子

有些马是纯种的。

即

x是一匹马，且**x**是纯种的。

然而，在这两个例子中，字母x的用法是不同的。在第一个例子中，我们想说的是，无论x可能是什么，即对于任何x，所断言的都为真。但在第二个例子中，我们想断言的只是对于某个x。在目前使用的符号体系中，对于任何被记作∀，对于某个被记作∃。于是。两个句子可以写为如下形式：

（∀**x**）（如果**x**是一匹马，那么**x**是一个哺乳动物），

（∃**x**）（**x**是一匹马，且**x**是纯种的）。

符号∀是一个倒写的A，暗示“所有”（All），称为全称量词。符号∃是一个向后写的E，旨在暗示“存在”（exists），称为存在量词。于是第二个句子可以读作

存在一个**x**，使得**x**是一匹马且**x**是纯种的。

通常把逻辑关系如果……，那么……记为⊃，把……且……记为∧。利用这些符号，上面的句子就变成：[9]

（∀**x**）（**x**是一匹马⊃**x**是一个哺乳动物），

（∃**x**）（**x**是一匹马∧**x**是纯种的）。

这可以被缩写为下面的形式：

（∀**x**）（马（**x**）⊃哺乳动物（**x**）），

（∃**x**）（马（**x**）∧纯种的（**x**））。

甚或更简洁地写成：

（∀**x**）（马（**x**）⊃哺（**x**）），

（∃**x**）（马（**x**）∧纯（**x**））。[16]

在上一章中，我们举了乔和苏珊用逻辑来寻找乔的支票簿的例子。在那个例子中，我们用字母简化句子如下：

L=乔把支票簿忘在了超市。

F=乔的支票簿在超市找到了。

W=乔昨晚在饭馆开了一张支票。

P=昨晚开了支票之后，乔把支票簿放在了他的夹克口袋里。

H=乔从昨晚起就没有再用过他的支票簿。

S=乔的支票簿仍然在他的夹克口袋里。

他们的推理可以归结为如下模式：

前提：

如果L，那么F。

非F。

W且P。

如果W并且P并且H，那么S。

H.

结论：

非L。

S.

如果用符号¬ 来表示“非”，并且利用我们已经介绍过的其他符号，则上面的句子就变成：

$$L \supset F$$

$$\neg F$$

$$W \wedge P$$

$$W \wedge P \wedge H \supset S$$

$$H$$

$$\neg L$$

$$S$$

还有最后一个符号：∨代表……或……。下表是对我们已经介绍过的符号的总结：

¬ 非……

∨ ……或……

∧ ……且……

⊃如果……，那么……

∀任何

∃某个

在上一章的结尾，我们举了一个逻辑结构无法用布尔的方法进行分析的例子，即

所有失败的学生或是糊涂的或是懒惰的。

而对弗雷格的逻辑而言，它是很简单的。如果用

F（**x**）表示**x**是一个失败的学生，

S（**x**）表示**x**是糊涂的，

L（**x**）表示**x**是懒惰的，

则这个句子可以表示为：

$$(\forall x)(F(x) \supset S(x) \vee L(x))。$$

现在我们应该很清楚了，弗雷格并不仅仅是对逻辑发展了一种数学处理，他实际上创立了一种新的语言。在这一点上，他以莱布尼茨的普遍语言思想为导向，只要恰当地选择符号，语言就可以获得力量。[10]这种语言的表达能力可以从下面的例子中体现出来，其中我们用L（x, y）来表示x爱y：

每个人都爱某人 $(\forall x)(\exists y)x$爱y $(\forall x)(\exists y)L(x, y)$

某人爱每个人 $(\exists x)(\forall y)x$爱y $(\exists x)(\forall y)L(x, y)$

每个人都被某人所爱 $(\forall y)(\exists x)x$爱y $(\forall y)(\exists x)L(x, y)$

某人被每个人所爱 $(\exists y)(\forall x)x$爱y $(\exists y)(\forall x)L(x, y)$

下面是另外一个例子：

每个人都爱一个情人。

我们先把它写成

$$(\forall x)(\forall y)[y\text{是一个情人} \supset L(x, y)]。$$

如果我们把“一个情人”理解为“爱某个人”，则我们就可以用（∃z）L（y, z）来代替y是一个情人，最后我们得到：

$$(\forall x)(\forall y)[(\exists z)L(y, z) \supset L(x, y)]。$$

弗雷格发明形式句法

布尔的逻辑只不过是需要用普通数学方法进行发展的另一数学分支，这当然包括使用逻辑推理。但是用逻辑来发展逻辑有些循环。在弗雷格看来，这是不可接受的。他的目标是要表明一切数学如何可能被建立在逻辑的基础之上。为了令人信服地做到这一点，弗雷格必须找到某种不用逻辑来发展他的逻辑的方法。他的解决办法是用冷酷无情的精确的语法规则或句法规则把他的概念文字发展成一种人工语言。这就使得把逻辑推理表示为纯粹机械运算即所谓的推理规则成为可能，这些规则仅仅与符号排列的样式有关。这也是第一次用精确的句法构造出形式化的人工语言。从这个观点看，概念文字是我们今天使用的所有计算机程序设计语言的前身。

弗雷格最基本的推理规则是这样来使用的：如果◇和△是用弗雷格概念文字写成的任意两个句子，那么如果◇和（◇⊃△）都被断言，则句子△也可以被断言。关于这一演算，我们只需注意它的执行，而不需要理解⊃是什么意思。当然，我们可以看出这条规则是不可能导致错误的，因为它仅仅能够使我们从◇和（如果◇，那么△）推出△。但在实际运用这条规则的时候，只需把组成◇这个句子的符号与较长句第一部分中的符号进行逐一对应。[11] 在那个寻找乔的支票簿的例子中，我们的前提是：

$$W \wedge P \wedge H \supset S。$$

如果我们也能够断言W∧P∧H，那么该规则就可以使我们断言所需的结论之一，即S。下面就是对应的情况：

$$W \wedge P \wedge H \supset S。$$

$$W \wedge P \wedge H。$$

弗雷格的逻辑已经成了在数学系、计算机科学系和哲学系教给本科生的标准逻辑。[12] 它曾是众多研究领域的基础，并且间接地帮助图灵表述了通用计算机的思想。但我们现在讲这些还为时过早。

弗雷格的逻辑比布尔的逻辑前进了一大步。第一次有一个精确的数理逻辑系统至少在原则上包含了数学家们通常使用的全部推理。但在达到这个目标的过程中，也有一些东西失去了。从弗雷格逻辑中的某些前提出发，我们可以尝试运用弗雷格的规则以获得希望得到的结论。但如果这一尝试失败了，弗雷格就没有办法知道，这到底是由于智力发挥得不够或坚持不够彻底，还是因为希望得到的结论根本就无法从前提中导出。这一缺陷意味着弗雷格的逻辑还没有实现莱布尼茨之梦，即说一句“让我们算一下”，那些知道逻辑规则的人就可以成功地判定某个结论是否可以导出。

伯特兰·罗素的信为何如此具有毁灭性

如果弗雷格的逻辑是一项这么了不起的成就，那么罗素的信为什么会令弗雷格如此沮丧呢？弗雷格认为他的逻辑只是朝着为算术提供完备的基础迈出了一步。尽管人们已经从莱布尼茨和牛顿的微积分中得出了丰硕的成果，但数学家们习惯于使用的某些推理步骤是否是正当的，这还很成问题。19世纪时，这些问题终于通过新发展出的一种深刻的数系理论而逐步得到了澄清。最终一切都归结为所谓的自然数：

1，2，3，……

弗雷格希望能为自然数提出一种纯粹逻辑的理论，从而证明算术、微积分的所有进展乃至一切数学都可以被看成逻辑的一个分支。这种后来被称为逻辑主义的观点也同样为伯特兰·罗素所持有。逻辑主义曾被美国逻辑学家阿隆佐·丘奇解释为：主张逻辑与数学的关系就是同一学科的基本部分和高等部分之间的关系。[17]

于是，弗雷格便希望能用纯逻辑术语来定义自然数，然后再用他的逻辑导出它们的性质。例如，3这个数将被解释为逻辑的一部分。这如何可能呢？自然数是集合的一种属性，即它的元素的数目。3这个数是为下面所有事物所共有的某种东西：三位一体、并驾拉车的三匹马的集

合、三叶草的叶子的集合、字母{a，b，c}的集合。显然，我们可以看到这些集合中任意两个的元素数目都相等。我们可以把它们一一对应起来。弗雷格的思想是要把3这个数等同于所有这些集合的集合。也就是说，3这个数只不过是所有3个一组的事物的集合。一般而言，一个给定集合的元素数目可以被定义为：能够与给定集合一一对应的所有那些集合的集合。[13]

弗雷格关于算术基础的两卷著作阐述了如何利用他的《概念文字》所提出的逻辑发展出自然数的算术。伯特兰·罗素在1902年的信中向弗雷格指出，他的整个工作是不一致的，也就是说是自相矛盾的。事实上，弗雷格的算术使用了集合的集合。罗素在信中指出，用集合的集合进行推理很容易导致矛盾。罗素的悖论可以这样来说明：如果一个集合是它自身的一个成员，那么就把这个集合称为异常的；否则就称它为正常的。一个集合如何可能是异常的？罗素自己举的关于异常集合的例子是：所有那些能够用数目少于19的英语单词定义的东西所组成的集合（the set of all those things that can be defined in fewer than 19 English words）。由于方才我们仅用16个（英文）词就把这个集合定义了，所以它属于它自身，从而是异常的。另一个例子是：所有不是麻雀的东西所组成的集合。无论这个集合是什么，它都肯定不是一只麻雀，所以这个集合也是异常的。

罗素指出，如果把所有正常集合所组成的集合记为ε，那么ε是正常的还是异常的？答案必定非此即彼。但情况似乎并非如此。ε是正常的吗？如果ε是正常的，那么由于ε是由所有正常集合所组成的集合，它就将属于自身。而这恰恰说明它是异常的。如此一来，ε就只能是异常的了。可是，由于它是由所有正常集合所组成的集合，所以它将不属于自身。而这恰恰又使它成为正常的！无论哪个结果都导致了矛盾！

罗素的悖论是大量令人大伤脑筋的有趣难题之一。而当弗雷格收到罗素的信时，他可没有觉得有趣。他马上意识到，这一矛盾可以从他用于发展算术的系统中导出。如果一则数学证明陷入矛盾，那么就证明该论证的前提之一是错误的。从古至今，这条原则都被当成一种有用的证明方法：要证明一个命题，我们只要说明否定它会导致矛盾就可以了。但是对于可怜的弗雷格来说，这一矛盾表明他的体系所基于的那些前提是靠不住的。弗雷格再也没有从这个打击中恢复过来。[14]

弗雷格和语言哲学

1892年，弗雷格在一份哲学杂志上发表了一篇论文，它的题目或可译为“论含义与指称”（*On Sense and Denotation* ）。[15] 除了弗雷格的逻辑，正是这篇论文的内容才使哲学家对他的工作如此感兴趣。

弗雷格指出，我们可以用不同的语词来命名或指称同一个具体对象，尽管它们有着非常不同的含义或意义。他的著名例子是“晨星”和“昏星”。它们的含义是非常不同的：一个是指日落后看到的星星，另一个是指日出前看到的星星。但两者都指称同一颗星体——金星。两者指的是同一对象，这个事实并不是显然的，它一度曾是一项真正的天文发现。弗雷格的某些考虑与替代性有关：考虑下面这个句子

金星是晨星，

它与

金星是金星

是非常不同的，尽管相对于其中的某一句话而言，另一句话只是替换了一个指称相同对象的语词而已。

这些思想标志着20世纪哲学的一个主要分支——语言哲学的开端。[16] 可以说，当代计算机科学的某些重要概念也是源自这篇文章。[17]

弗雷格与莱布尼茨之梦

弗雷格认为他的《概念文字》体现了莱布尼茨所憧憬的逻辑的普遍语言。的确，弗雷格的逻辑可以处理林林总总、各不相同的学科。但在莱布尼茨看来，这很可能会令他失望。它至少在两个方面未能如他心意。莱布尼茨曾经设想过这样一种语言，它不仅能够进行逻辑演绎，而且也能自动包含科学与哲学中的一切真理。只有面对着基于实验和理论化的科学在18至19世纪突飞猛进，这种天真的期盼才是可以设想的。

从我们的主题来看，指出弗雷格逻辑的另一个局限性是更恰当的。莱布尼茨曾经设想过一种能够成为有效的演算工具的语言，这种语言将

通过对符号的直接操作而使逻辑推理自动进行。事实上，在弗雷格的逻辑中，除了那些最简单的演绎，其余的情况都复杂得让人难以忍受。这些演绎过程不仅长得可怕，而且在弗雷格概念文字的逻辑中，他的规则也没有为判定某个结论是否可以从给定前提中推导出来而提供演算步骤。

由于概念文字的确完全包含了普通数学中所用到的逻辑，所以用数学方法来研究数学活动本身也就成为可能了。正如我们将会看到的，从这些研究中得出了一些让人意想不到的非凡成果。能否找到一种计算方法，它能够说明在弗雷格的逻辑中某一推理是否是正确的呢？这种探究在1936年达到了高潮，其结果是这样一则证明：没有这样的一般方法存在。对于莱布尼茨的梦想来说，这并不是一条好消息。然而，正是在证明这条否定性结论的过程中，阿兰·图灵发现了某种必定会令莱布尼茨欣喜的东西：他发现原则上可能设计出一种通用机器，它能够执行任何可能的计算。

第4章　康托尔：在无限中摸索

1，2，3，……的序列永无尽头，这就是所谓的自然数或计算数。无论开始的数有多大，你总能通过加上1来得到一个更大的数。有人也许会认为自然数是通过一个过程产生出来的，即从1开始连续不断地加1：

1，1+1=2，2+1=3，……，99+1=100，……

这样一个不断超越任何有限界限的过程，被亚里士多德称之为一种潜无限。然而，亚里士多德不愿认为这个过程的终点——所有自然数的无限集——是合理的。这将成为一种“完成的”无限或“实”无限，亚里士多德宣称这是不合理的。[1] 亚里士多德的观点极大地影响了12世纪的经院哲学家，特别是托马斯·阿奎那。无限的本性问题一直在困扰着数学家、哲学家和神学家。神学家们声称，一种完成了的无限实际上是上帝的一个特性，他们断言，对于人来说，它只能永远是一个秘密。这样的说法并没有使莱布尼茨感到气馁，他写道：

> 我是如此地赞同实无限，以至于我不但不像通常所说的那样承认自然厌恶它，而且还认为自然无处不在频繁地利用它，为的是更加有效地显示造物主的完美性。[2]

在18和19世纪对于数学变得如此重要的微积分的极限过程正是潜无限的例子。关于这一点，伟大的德国数学家卡尔·弗里德里希·高斯（1777—1855）曾经警告说：

> 我极力反对把无限当成一种完成的东西来使用，这在数学上是绝不能允许的。无限只不过是一种言说方式，使人可正确地谈论极限罢了。[3]

19世纪中叶以后，从当时人们所关注的问题中自然产生出来的数学问题似乎要求在其精确表述中使用完成了的无限。在那些解决这一问题的数学家当中，只有格奥尔格·康托尔不顾高斯的警告，迎接挑战去创

立一种关于实无限的深刻而融贯一致的数学理论。康托尔的工作引发了一阵阵的批评浪潮：不仅是数学家，而且连哲学家和神学家也纷纷抨击这个人的鲁莽无礼，他竟然把数学科学的方法带进迄今为止神圣不可侵犯的无限领地。弗雷格支持康托尔对实无限的信念，他认识到了它对数学未来的重要性。同时，弗雷格看得很清楚，一场激烈的斗争就要在那些拥护和诅咒康托尔的无限的数学家之间展开了：

> 由于无限终将不会从算术中被排挤出去……于是我们可以预见，这个问题将会为一场关系重大的、决定性的斗争搭好舞台。[4]

当弗雷格写下这些话时，他不可能预见到，他本人所发展的算术基础将会成为这场斗争早期的一个牺牲品，成为伯特兰·罗素在10年之后的那封著名的信中的那个悖论的受害者，而罗素正是在研究康托尔的无限的过程中发现这个悖论的。弗雷格肯定想象不到，由康托尔的无限所引发的激烈的讨论、研究和争辩有一天会为通用数字计算机的发展提供重要的启发。

工程师还是数学家

1845年，格奥尔格·康托尔出生于俄国的圣彼得堡，似乎没有任何迹象显示他日后会在德国大学任数学教授。康托尔的母亲玛丽·伯姆出生于一个著名的音乐家庭，她本人也是一位颇有成就的音乐家。他的父亲格奥尔格·瓦尔德玛·康托尔出生于哥本哈根，但他小时候被带到了圣彼得堡。据说在那里，他是在一家路德新教的慈善机构中被抚养长大并接受教育的。虽然玛丽已经受过洗而成了一名罗马天主教徒，但在结婚以后改信了路德宗，格奥尔格·康托尔和他的3个兄弟姐妹就是在这种信仰中被抚养长大的。[5]

格奥尔格·瓦尔德玛·康托尔是一个非常成功的商人。他起先是圣彼得堡的一个批发商，后来又成了圣彼得堡证券交易所的经纪人。一位作者在谈到康托尔的父亲给还在上学的康托尔写的信时，曾动情地说：

> 这个既多才多艺又有教养、既成熟又和蔼的人简直让人着迷。在它们[这些信件]的字里行间透出一种通常在成功的商人那里找不

到的精神。[6]



格奥尔格·康托尔

不仅穷人特别难逃结核病这个19世纪的巨大灾祸的厄运，连富人也很难幸免。康托尔的父亲就是染上了这种可怕的疾病而最终宣告不治。虽然格奥尔格·瓦尔德玛此时才40多岁，但疾病却使他不得不终止自己的生意，在儿子11岁时举家迁往德国。不过他是如此地成功，以至于在他搬迁和去世7年之后，他的4个孩子仍然能够获得很好的抚养。

格奥尔格·瓦尔德玛相信工程师的职业最适合发挥他儿子的天分，但是让格奥尔格大为高兴的是，他最终默许了这个孩子当一名数学家的愿望。在柏林，年轻的格奥尔格·康托尔有幸跟3位大数学家学习，他们是：卡尔·外尔施特拉斯、恩斯特·库默尔和利奥波德·克罗内克。康托尔的数学兴趣一开始集中在非常传统的领域。在他的职业刚刚开始时，很难设想他竟然注定会沿着革命性的方向拓展数学思想的疆域，他的老师克罗内克也将成为他强硬的对手，把他一生的工作斥之为无稽之谈。

从弗雷格的故乡耶拿沿萨勒河往上游走35英里就到了工业城市哈勒，正是在那里，康托尔得到了他第一个大学教职，并将度过自己的后半生。与当时德国学术生涯开始的典型情况一样，康托尔也是被任命为一名无俸的编外讲师。在这种情况下，独立的经济来源显然对于开展学术事业来说是必需的。哈勒大学的领衔数学家爱德华·海涅发现了康托尔超凡的数学能力，并建议他研究一些与无穷级数有关的问题。在第一章中，我们碰到了无穷级数，即莱布尼茨著名的

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\frac{1}{11}+\cdots$$

在这样的级数中遇到的“无穷”只是潜无穷，它就是高斯头脑中想到的那种类型。对于一个无穷级数来说，随着一项项地加上去，我们离极限值越来越近了（对于莱布尼茨级数来说，这个极限是π/4），我们说级数收敛于这个极限。一个完成了的无限是不成问题的；在这个过程中的任何一个阶段，我们都只是加上了有限多个数。

很自然地，无穷级数的问题在自莱布尼茨时代以来的两个世纪里已

经有了长足的发展。康托尔研究了三角级数 [7] （之所以这样称呼，是因为级数中的项包含了三角学中的正弦和余弦）。他希望弄清楚在何种情况下两个这种类型的不同级数会收敛于同一值，事实上是希望证明这样的情况是非常罕见的。这项研究把康托尔远远带出了原先的研究领域：他发现为了得到希望的结果，他不得不把无穷集当作完成了的整体来处理，并且对其进行复杂的运算。不久，他就把集合论（*Mengenlehre* ）发展成了一门独立的学科。

无穷集的大小是不同的

即使我们认为把所有自然数1，2，3，……所组成的集合当成一个完成了的实无限来处理是有意义的，如果我们问这个集合中有多少个数，这难道也是有意义的吗？是否有无限大数可以用来数无限集呢？不反对这种完成的无限的莱布尼茨在一封致天主教神父、神学家和哲学家尼古拉·马勒伯朗士的信中讨论了这个问题。他的结论是，这种无限大的数并不存在。他的推理可以解释如下：只要一个集合中的元素可以同另一个集合中的元素一一对应起来，我们就可以说这两个集合的元素数目是相等的，即使我们不知道这个数目到底是多少 [18] 。例如，如果我们发现一个礼堂中既没有空座位，又没有站着的人，那么我们（不必清点）就可以下结论说，礼堂中的人数与座位数是相等的——我们其实是把每一个座位与每一个人一一对应了起来。莱布尼茨认为，如果真有无限数这样的东西存在，那么同样的思想也可以应用于它们：倘若两个无限集之间可以建立起一个一一对应的关系，则我们就可以下结论说，这两个集合拥有同样数目的元素。接着，他建议把这一观念运用到如下两个集合上：由所有自然数所组成的集合1，2，3，……以及由所有偶数所组成的集合2，4，6，……在这两个集合之间很容易设计出一种一一对应的关系，我们只要把每一个自然数与它的二倍对应起来就可以了：

$$
\begin{matrix} 1 & 2 & 3 & 4 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 2 & 4 & 6 & 8 & \cdots \end{matrix}
$$

请注意，即使自然数集与偶数集都是无限集，在这两个集合之间也可以建立起这种一一对应关系。比如对应着117这个自然数的偶数是234，对应着4228这个自然数的偶数是8456等等。莱布尼茨推论说，如果真有像无限数这样的东西，那么这种一一对应的存在就会迫使我们下结论说，自然数的数目与偶数的数目是相等的。但这怎么可能呢？自然数中并非只有偶数，而且还有所有的奇数，它们又组成了一个无限集。有一条最基本的数学原理就是整体大于它的任何部分，这可以追溯到欧几里得。[8] 于是，莱布尼茨得出结论说，所有自然数的数目这一概念是不协调的，谈论一个无限集中元素的数目是没有意义的。正如他所说：

> 对于任何一个数，都存在着一个与之对应的偶数，那就是它的二倍。因此所有数的数目并不比偶数的数目更多，也就是说，整体并不比部分大。[9]

康托尔也做了和莱布尼茨一样的推理，从而面临着同样的困境：或者谈论一个无限集中元素的数目是没有意义的，或者某些无限集将与它的一个子集具有相同的元素数目。然而，莱布尼茨选择了前者，康托尔却选择了后者。他开始发展一种能够应用于无限集的关于数的理论，这种理论恰恰认为一个无限集将与它的一个部分拥有同样多的元素数目。

沿着莱布尼茨没有走完的道路，康托尔开始研究在两个不同的无限集之间建立一一对应在什么情况下是可能的。尽管莱布尼茨已经发现在自然数集与它的一个子集（偶数集）之间可以建立起一种一一对应关系，但康托尔还考虑了似乎比自然数集更大的集合。他所考虑的一个例子是可以被表示为（正）分数的数的集合，[10] 比如1/2或5/3。由于自然数可以用分母为1的分数来表示（比如7/1），所以自然数集可以被看成这个集合的一个子集。但是，康托尔稍加考虑，便发现他可以在由这些分数所组成的集合与自然数集之间建立起一种一一对应关系。分数可以像下面这样排成一个序列：

$$\left|\ \frac{1}{1}\ \middle|\ \frac{1}{2}\ \frac{2}{1}\ \middle|\ \frac{1}{3}\ \frac{2}{2}\ \frac{3}{1}\ \middle|\ \frac{1}{4}\ \frac{2}{3}\ \frac{3}{2}\ \frac{4}{1}\ \middle|\ \frac{1}{5}\ \frac{2}{4}\ \frac{3}{3}\ \frac{4}{2}\ \frac{5}{1}\ \right|\cdots$$

它们是按照每一个分数的分子与分母之和进行分组的：首先是和为2的分数（这样的分数只有1个），然后是和为3的分数（这样的分数有2个），然后是和为4的分数（这样的分数有3个），再后是和为5的分数（这样的分数有4个），依此类推。现在很容易在它们与自然数之间建立起一一对应的关系：

$$\begin{array}{ccccccccccccccc} \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{2}{2} & \frac{3}{1} & \frac{1}{4} & \frac{2}{3} & \frac{3}{2} & \frac{4}{1} & \frac{1}{5} & \frac{2}{4} & \frac{3}{3} & \frac{4}{2} & \frac{5}{1}\cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15\cdots \end{array}$$

由于从直观上看，分数要比自然数多许多，所以这一证明可能会使我们猜想，每一个无限集都可以与自然数之间建立起一一对应的关系。康托尔的伟大贡献就是说明了情况并非如此。可以用分数来表示的数被称为有理数。如果一个有理数是用小数来表示的，那么最终将会有重复的数字出现。下面是一些例子：

$$\frac{1}{3} = 0.3333333333333333333333\cdots$$

$$\frac{1}{4} = 0.2500000000000000000000\cdots$$

$$\frac{5}{3} = 1.6666666666666666666666\cdots$$

$$\frac{24}{11} = 2.1818181818181818181818\cdots$$

$$\frac{9}{7} = 1.2857142857142857\cdots$$

可以用小数来表示的数被称为实数，无论它们最终是否会有数字重复。其小数表示不会重复的那些数被称为无理数。下面是一些无理数的例子：

$$\sqrt{2} = 1.41421356237309505 0\cdots$$

$$\sqrt[3]{2} = 1.25992104989487316 0\cdots$$

$$\pi = 3.14159265358979324 0\cdots$$

$$2^{\sqrt{2}} = 2.66514414269022519 0\cdots$$

像$\sqrt{2}$和$\sqrt[3]{2}$ 这样的数连同所有的有理数被称为代数数，这是因为它们可以作为代数方程的根。（$\sqrt{2}$ 是方程$x^2=2$的一个根，$\sqrt[3]{2}$ 则是方程$x^3=2$的一个根）。而业已证明，π和$2^{\sqrt{2}}$ 不可能是任何代数方程的根，这样的数被称为超越数。

在表明了分数可以与自然数之间建立起一一对应的关系之后，康托尔把注意力转向了由所有代数数所组成的集合，他又一次轻松地找到了一种方法，可以把它们与自然数一一对应起来。很自然地，他想知道这是否对于由所有实数所组成的集合来说也是成立的。我们可以看看28岁的康托尔在1873年给理查德·戴德金（康托尔于前一年在瑞士休假时很偶然遇到的一位年轻数学家）写的信中的思考过程是怎样的。康托尔那时刚刚晋升为哈勒大学的数学教授，他致信戴德金说，（正如我们已经看到的那样）我们可以在自然数与包含更广的所有正分数之间构造出一种一一对应的关系。他甚至说明了这个结论对于由所有的代数数所组成的集合也是正确的。在信中，康托尔提出了自然数集与实数集之间是否可以建立起一一对应的问题。戴德金的回复表明，他对这个问题没有什么兴趣。大约一周之后，康托尔在另一封信里就已经能够向戴德金证明这个非凡的结论了，即实数集无法与自然数集之间建立起一一对应，无限集至少有两种大小。

显然，康托尔本人并不十分确定这一发现是否值得发表。只是在他以前的老师卡尔·外尔施特拉斯鼓励他之后，他才把它拿出去发表。康托尔的工作的革命性含义在那篇四页的论文中表现得并不明显。这篇论文的重点并不是无穷集的大小不只一种这一事实，而是对实数中存在着超越数的一种新的证明，后者只是前者的一个推论。康托尔的证明的大意是，由于代数数可以与自然数一一对应起来，而实数则无法进行这种对应，所以实数集与代数数集是不同的。因此就必定存在着一个实数不是代数数，于是它就是超越数。[11]

与此同时，康托尔的个人生活也丰富了起来。1874年，他与瓦里·古特曼（Vally Guttman）结了婚，古特曼是他妹妹的一个好朋友，也是一个很有天分的音乐家。他们后来有了六个孩子，从任何方面来说，这都是一个情意绵绵、温暖如春的家庭。尽管康托尔在职业圈子里有强硬甚至是难于相处的名声，但他在家里显然是温文尔雅的。有人是这样描述康托尔一家吃饭时的情形的：

> 在吃饭的时候，他会静静地坐着，让他的孩子们挑起话题，然后起身感谢他的妻子做了这顿饭，并说："你对我满意吗？你还爱我吗？"[12]

然而，随着他把越来越多的精力投入到集合论的研究中，康托尔那令人不安的新思想开始遭到越来越多的人的反对。他以前的老师克罗内克后来成了康托尔整个研究方向的不遗余力的反对者，他甚至还阻止他发表某些论文。在这种气氛下，康托尔找不到一所大学给他职位，在那里他能遇到与他具有相同水平的同事，他不得不仍然待在哈勒这个死气沉沉的地方。甚至连康托尔劝说他的朋友戴德金来哈勒的努力后来也失败了。1886年，康托尔听从了命运的安排，他在哈勒买了一所漂亮的房子和家人住了下来。

康托尔对无限数的探求

虽然高斯警告数学家不应与完成的无限打交道，但康托尔顾不了这些，他感觉到自己被无限的诱饵牢牢地套住了。迄今为止，无限一直是神学家和形而上学家的领地。他的数学研究已经为他那激进的思想提供了基础，但他大大超越了那些研究所规定的范围。在日常语言中，人们

用两种不同但却相互有关联的方式来使用自然数1，2，3，……它们分别被用于计数和排序，比如下面这两个句子：

·这个屋子里有四个人。

·乔的马得了第四名。

日常语言用基数和序数之间的差别来体现这一点：用基数一、二、三……；序数用第一、第二、第三……基数用于指明某个集合中有多少个东西，序数则用于指明这些东西是如何以一定次序排列的。康托尔关于在自然数与实数之间不存在一一对应的发现促使他思考了无穷基数，他关于三角级数的工作则暗示了一种把无穷序数概念化的方法。

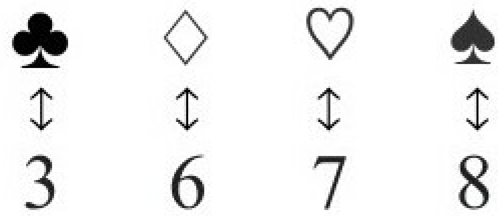
康托尔假定每一个集合（无论是有限的还是无限的）都有唯一一个基数。康托尔认为一个集合的基数可以这样来得到：丢掉集合元素的特殊性质，剩下的就是毫无特征的单元。特别地，如果两个集合可以被一一对应起来，那么它们就有同一个基数。设M表示某个完全任意的集合，康托尔引入记号$\overline{\overline{M}}$表示集合M的基数。[13] 例如，[19] 如果

$$A=\{\clubsuit\ \heartsuit\ \diamondsuit\ \spadesuit\},B=\{3,6,7,8\},\text{且}\ C=\{6,5\},$$

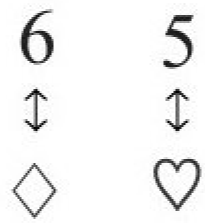
那么

$$\overline{\overline{A}}=\overline{\overline{B}}=4,\ \overline{\overline{C}}=2。$$

当然，A与B之间很容易建立起一个一一对应关系：



如果两个集合基数不同，那么会发生什么情况？用符号来讲，这个问题就是说，对于集合M与集合N而言，什么时候$\overline{\overline{M}} \neq \overline{\overline{N}}$。在这种情况下，两个基数中一个较大一个较小。利用标准符号＜（“小于”）和＞（“大于”），我们可以$\overline{\overline{M}} < \overline{\overline{N}}$（或者等价的$\overline{\overline{M}} > \overline{\overline{N}}$）来表示N具有较大的基数。为了证明这一点，我们只需在M和N的某个子集之间建立一个一一对应关系。[14] 于是，在上面那个$\overline{\overline{A}} > \overline{\overline{C}}$例子中（因为4＞2），A的子集{（◇♡）}可以这样与C对应：



只要限于有限集，那么所有这些似乎都只不过是在用晦涩的术语表示简单而熟悉的事物。的确，康托尔思想的力量只有在运用到无限集时才能表现出来。康托尔把无限集的基数称为超限数。他关于超限数的第一个例子是自然数集的基数，并用符号$\aleph_0$来表示。$\aleph_0$通常读作“阿列夫零”，$\aleph$是希伯莱字母表上的第一个字母。[20]

康托尔用符号来表示实数集的基数（因为实数集有时被称为连续统[continuum]）。康托尔深信，C就是继$\aleph_0$之后的下一个超限基数。$\aleph_0$与C之间不存在别的基数，这一猜想被称为连续统假设。尽管康托尔为证明这个猜想付出了多年艰辛的努力，但最终未能解决这个问题：他既不能证明连续统假设为真，也不能证明连续统假设为假。这一失败带给了康托尔无尽的痛苦。今天我们知道，可怜的康托尔是在枉费心机。库尔特·哥德尔于1938年和保罗·科恩于1963年的重大发现揭示了，如果

连续统假设问题可以被解决，那么就必须超越普通数学的方法。所以康托尔没能解决这个问题也就不足为奇了。事实上，直到今天，哥德尔—科恩的否定结果是否就是最好的结果，以及是否可能有新的强有力的方法能够得出一个更加令人满意的结果，专家们对此仍然莫衷一是。

在康托尔关于三角级数的工作中，他考虑了某种可以被一步步反复应用的特定过程：第一步、第二步、第三步等等。然而使康托尔跨入超限领域的是，他意识到在所有这些无限步之后还有更多的步。不久，他开始谈及第ω步，第（ω+1）步以及更多的步，并且发展出后来被他称为超限序数的算术。[21] 让我们首先看一看有限集{♣◊♡}。其成员可以按照六种不同的方式加以排列：

1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
♣	◊	♡	♣	♡	◊	◊	♣	♡	◊	♡	♣	♡	♣	◊	♡	◊	♣

然而，这6种排列都显示出同一种样式：第一项后面跟着第二项，第二项后面跟着第三项。任何有限集都是这样的：对其成员进行排列的所有不同方式的背后都显示出同一种样式。如果一个集合有n项，那么任何排列都会显示第一项、第二项……最后是第n项。康托尔发现无穷集的情况是完全不同的。无穷集可以以不同的方式排列成非常不同的样式。例如，假定把自然数1，2，3，……排列成所有偶数成员在所有奇数成员之前：

2，4，6，……1，3，5，……

如果我们试图用序数来说明每一项在序列中的次序，那么我们就会发现，我们所熟悉的有限的序数都被偶数用尽了：

$$
\begin{array}{cccccccc}
1^{st} & 2^{nd} & 3^{rd} & \cdots & ? & ? & ? & \cdots \\
\downarrow & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\
2 & 4 & 6 & \cdots & 1 & 3 & 5 & \cdots
\end{array}
$$

康托尔发现了如何用超限序数来解决这个困难。于是，在所有的有限序数之后，康托尔又设定了第一个超限序数，他用希腊字母ω来表示，在它之后是ω+1，ω+2等等。然后康托尔就可以很容易地为上例中的奇数排序了：

$$
\begin{array}{cccccccc}
1^{st} & 2^{nd} & 3^{rd} & \cdots & \omega^{th} & (\omega+1)^{th} & (\omega+2)^{th} & \cdots \\
\downarrow & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \\
2 & 4 & 6 & \cdots & 1 & 3 & 5 & \cdots
\end{array}
$$

康托尔发现，自然数可以用越来越大的序数以许多不同的方式加以排列。他称有限序数即自然数1，2，3，……为第一数类，称为自然数的不同安排排序的超限序数为第二数类。康托尔在研究构成第二数类的超限序数集时，用符号 $\aleph_1$ 来表示它的基数。值得注意的是，康托尔能够证明， $\aleph_1$ 就是紧接在最小的超限基数$\aleph_0$之后的下一个基数。因此，$\aleph_1 > \aleph_0$，，不存在大于$\aleph_0$而小于$\aleph_1$的基数。

康托尔一直努力试图证明的连续统假设是，C 就是继$\aleph_0$ 之后的下一个基数。他已经知道，$\aleph_1$ 就是$\aleph_0$之后的下一个基数，所以连续统假设可写成这样一个简洁的问题：

$$C \overset{?}{=} \aleph_1。$$

不幸的是，康托尔虽然写下了这个方程，但他并没有证明它为真。

既然有第一数类和第二数类，那么是否有第三数类呢？绝对有！在为基数为$\aleph_1$的集合排序时，第一数类和第二数类的数就不够了。康托尔把第三数类最开始的超限基数记为ω_1，并把第三数类中的所有序数所组成的集合的基数称为$\aleph_2$。然后，$\aleph_2$就是紧接在$\aleph_1$之后的下一个基数。康托尔发现这个过程是永无尽头的：$\aleph_2$之后还有$\aleph_3$，$\aleph_3$之后还有$\aleph_4$等等。在所有这些之后是$\aleph_\omega$等等。

当康托尔提出这些思想时，他实际上是在对一个从未有人到过的领域进行着探索。他无法依靠任何数学规则，而不得不凭借自己的直觉独自创造。考虑到他所研究的领域的本性，他的工作一直顺利地进行了下去，这是令人赞叹的。但是从一开始，反对康托尔整个事业的呼声就此起彼伏。克罗内克的反对我们已经提到了。有一则在数学家之间流传很广的故事说，著名的法国数学家昂利·庞加莱曾经说，总有一天，康托尔的集合论“会被看作一种被征服了的疾病”。尽管这个故事似乎是伪造的，但从它的流行程度就可看出康托尔那时正面临着怎样的挑战。

对角线方法

如果说今天的学生从康托尔的成果中学到了某种东西，那么几乎可以肯定这就是他的对角线方法。这种方法是康托尔于1891年在一篇仅有四页的论文中发表的，这时他几乎已经停止了数学研究，他那些讨论超限数的重要文章也已经发表甚至重印。1874年，康托尔发表了他对自然数与实数之间不存在一一对应的证明，或者用他的符号表示，就是$\aleph_0<C$。这个证明借用了外尔施特拉斯所发展的极限过程的基本理论。利用对角线方法，同一结论可以从基本逻辑原理中得出。在我们的故事中，对角线方法将会不断出现。

在解释对角线方法时，使用一个贴着标签的包裹的比喻将是有益的。特别之处在于，被用作标签的东西就是包裹里的东西。举例说来，考虑一副纸牌中的4种花色：♣ ♡ ♢ ♠。让我们把其中的每一种花色用作一个“包裹”上的标签，而包裹中所包含的也正是这些花色中的某些种，比如：

♣	◇	♡	♠
↕	↕	↕	↕
{ ◇ ♡ }	{ ◇ ♠ }	{ ◇ ♡ ♠ }	{ ♣ ◇ }

我们可以用表格的形式来表示这一信息，其中+表示某一项在包裹内，—表示它不在包裹内：

	♣	◇	♡	♠
♣	⊖	+	+	−
◇	−	⊕	−	+
♡	−	+	⊕	+
♠	+	+	−	⊖

在这张表中，左边的一列是4种标签，包裹中的内容显示于各行中。对角线上的+和—外面围有圆圈，以示强调。对角线方法是一种把相同类型的项合到一个新包裹中去的技巧，它可以使这个新包裹中所含的东西与所有贴了标签的包裹都不同。它的工作过程是这样的：我们制作一张新的表格，并把相反的符号插到其对角线上。于是，由于与♣相对的是—，所以在我们的新表中，♣对应的是+。同样，♡对应—，◇对应—，♠对应+。结果如下：

	♣	◊	♡	♠
	+	−	−	+

于是，我们的新包裹就是{♣♠}。我们何以能够确信它与任何贴了标签的包裹都不同呢？情况是这样的，它不可能是标签为♣的包裹，因为♣不在那个包裹中，而在我们新的包裹中。它也不可能是标签为◊的包裹，因为◊在那个包裹中，而不在我们新的包裹中。依此类推。

现在，包裹当然指的就是集合，而标签则是一种在集合与其成员之间建立起一一对应的方法。这种方法是完全一般的：开始时的集合是否是一个无限集，这是无关紧要的。如果用那个集合中的每一个元素来做某个由同一种元素中的某些元素所组成的集合的标签，那么就可以用对角线方法来获得一个由那些元素所组成的新的集合，它与所有已经贴过标签的集合都不同。

让我们看看，如果我们从自然数集1，2，3，……开始，这种方法将怎样运作。设想把这些数中的某一些装进一个包裹。它可能只是由{7，11，17}组成的，还可能由所有偶数组成。现在，让我们设想把自然数本身作为标签，就像下面这个无穷序列一样：

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & \cdots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ M_1 & M_2 & M_3 & M_4 & \cdots \end{array}$$

其中M_1，M_2，M_3，M_4，……中的每一个都是由自然数所组成的包裹。这时，我们用下表来获得一个新的集合M，它不同于每一个集合：

1	一如果1在M_1中	否则+
2	一如果2在M_2中	否则+
3	一如果3在M_3中	否则+
4	一如果4在M_4中	否则+
…	… …	…

换句话说，如果1不属于M_1，那么1就属于M；如果2不属于M_2，那么2就属于M……依此类推。因此，M是一个不同于M_1，M_2，……的自然数集。由于M_1，M_2，M_3，M_4，……表示1，2，3，……与自然数集之间任何可能的一一对应，我们发现，没有什么对应可以包含自然数集的一切子集。换句话说，由一切自然数集所组成的集合即一切自然数的子集的集合的基数要大于$\aleph_0$。事实上，我们可以证明这个基数就是实数集的基数C。[15] 于是，对角线方法就为我们提供了另一种说明实数比自然数更多的方法。

这种方法非常具有一般性，它提供了另一种产生许多（与康托尔相继的$\aleph_s$不同的）超限基数的方法。例如，我们可以设想以实数作为标签的实数包裹。对角线方法表明，没有这样的标签可以包含一切实数集。因此，由所有这样的集合所组成的集合的基数必定大于实数集的基数C。[16] 而且，没有必要就此止步。直到今天，以这种方式获得的基数是如何与康托尔的$\aleph_0$，$\aleph_1$，$\aleph_2$，……纠缠在一起的问题仍然是困难和争论的一个来源。

沮丧和悲剧

从一开始，康托尔就面临着反对意见，他们主要反对这样一种观念，即生活在一个有限世界的有限的人类竟然期望能够对无限做出有意义的断言。不过就在世纪之交的时候，人们发现用康托尔的超限数进行不加限制的推理会导致非常荒谬可笑的结果，于是情况变得更加糟糕了。麻烦均源于把康托尔的超限基数或序数都收在单一集合中的尝试。如果存在着一个由所有基数所组成的集合，那么它的基数该是多少呢？它必须要比任何基数都大。但这又怎么可能呢？一个基数怎么可能比所有基数都大呢？

就在康托尔意识到这个令人不安的悖论之后不久，意大利数学家布拉里—福蒂在试图处理由所有超限序数所组成的集合时又发现了一个类似的困难：他表明，这样一个集合将导致一个比任何超限序数都大的超限序数，这显然是一个荒谬的结论。然后伯特兰·罗素登场了，他对所有这些给出了最为致命的一击。他考虑了这样一个问题：是否存在一个所有集合的集合？如果存在着这样一个集合，那么倘若把对角线方法应用于它，会出现什么结果？换句话说，如果我们考虑把任意多的集合打成包裹，然后用集合去给这些包裹贴标签，那么会出现什么情况？当然，我们将得到一个不同于所有那些已经拥有标签的集合的集合。正是在考虑这种情况时，罗素发现了他那个关于由一切不是自身成员的集合所组成的集合的著名悖论。这就是我们在上一章中谈到的罗素向垮掉的弗雷格所传达的悖论。

尽管罗素是通过思考康托尔的思想而发现他的悖论的，但这个悖论本身丝毫不依赖于有关超限数的考虑。在许多数学家看来，最基本的逻辑推理似乎已经变得不可靠了，在这当中充满了陷阱。毫不奇怪，大多数数学家仍然继续着自己通常的工作，他们离这些问题还很遥远。但是对于那些关心有关数学本性的基本问题的人来说，这种状况不啻为数学基础的一场危机。这些数学家和哲学家不久就发现他们分成了对立的阵营。特别是，一些人认为集合论是数学所必需的一部分，为此可付出任何代价，另一些人则力图使数学免于受到康托尔的超限数的污染。逻辑学家在20世纪头30年的工作就是被这些问题所主导的。

1884年，康托尔首次遭遇一系列精神崩溃，严重的沮丧状态持续了大约2个月。康复之后，康托尔把自己的精神问题归因于他关于连续统

假设的紧张工作以及克罗内克的攻击上。这个时候，他甚至给克罗内克写了一封信建议恢复友谊，克罗内克诚挚地回复了这封信。除了几件现在被认为是躁狂抑郁症的事件以外，康托尔对他的遭遇的解释在很多年里都被人广泛接受。现在一般认为，不管外部事件有多么严重，其精神紊乱的根本原因在于有缺陷的脑的特质，以及像康托尔与克罗内克意见不和这样的环境因素，连续统假设只是促成了而不是造成了主要的混乱。[17]

除了那篇已经提到的论述对角线方法的论文，这一事件大致标志着康托尔在集合论方面的奠基性工作的结束。在严重的精神疾病的间歇，康托尔研究了哲学、神学特别是莎士比亚戏剧的作者身份问题。对于康托尔来说，1899年是危机四伏、悲剧重重的一年。正是在这一年，他第一次遭遇了集合论悖论，他13岁的爱子之死也令他痛不欲生。

对学科只满足于懂得一点皮毛，这不是格奥尔格·康托尔的风格。他使自己成了一名研究伊丽莎白时期特别是莎士比亚戏剧的专家，他发表了一系列专著声称证明了这些剧本其实是出自弗兰西斯·培根之手。当然，这与集合论或超限毫无干系。不过，康托尔发现自己在哲学和神学方面的研究绝对与他关于无限的工作有关。康托尔相信，在超限之外还存在着一个绝对的无限，它仅靠人类的理解力是永远无法完全企及的。甚至是出自集合论的令人苦恼的悖论也应当从这一观点来理解。例如，所有超限基数的乘积应当被认为是绝对无限的，这就是为什么仅仅认为它们是超限的就会导致矛盾的原因。

一场决定性的战斗？

在德国哲学思想中，伊曼努尔·康德是一个至关重要的人物，他的批判哲学建立在两个关键问题之上：

·纯粹数学如何可能？

·纯粹自然科学如何可能？

康德对第一个问题的回答依赖于他所谓的对空间（对应于几何学）和时间（对应于算术）的“纯直观”。他认为这些直观完全独立于经验感觉。[18] 尽管康德强调了科学的重要性，但19世纪德国的后康德哲学却

沿着一条不同的方向发展了，它转向了一种认为观念和概念是第一位的绝对唯心论，认为世界似乎就是由它们构成的。这场运动的领导者之一是格奥尔格·弗里德里希·黑格尔，数以百计的忠实弟子都听过他的讲演。黑格尔拥有许多追随者（其中最著名的是卡尔·马克思和弗里德里希·恩格斯），直到今天，学者们仍然能够在他的著作中找到许多有价值的东西。然而，他有时做出的一些怪异推理只会招致嘲笑，特别是在他那部两卷本的《逻辑学》中，读者曾被要求对如下深刻的思想进行思考：

> 无等同于它自身。
>
> 存在即无。
>
> 无即存在。

这两个范畴在彼此转化的过程中融合为一个更进一步的范畴：生成。

与此同时，在那个世纪行将结束之时，部分是由于奥古斯特·孔德的“实证主义”思想，部分是由于科学的发展，一种新的“经验论”哲学开始在德国发展起来。在经验论者看来，理解世界主要依靠的是感觉材料。康托尔把这种经验论看作是对黑格尔式的胡说的一种反抗，但却发现它既粗糙又浅薄。经验论哲学的主要支持者之一，大科学家赫尔曼·冯·亥姆霍茨希望重新像康德那样把经验科学放到核心位置。他的一本论述计数和测量的小册子激怒了康托尔。1887年，康托尔在一篇从数学、哲学和神学观点讨论超限数的文章中，批评这本小册子表达了一种“极端经验心理学的观点，连带一种靠不住的教条主义”。他又接着抱怨说：

> 于是，在今天的德国，作为一种对华而不实的康德—费希特—黑格尔—谢林唯心论的回应，我们看到一种学院的实证主义的怀疑论正强有力地占据着统治地位。这种怀疑论甚至已经不可避免地把范围延伸到了算术，并且在这个领域导致了它最为致命的结论。这也许最终会对这种实证主义的怀疑论本身产生最大的毁灭作用。[19]

这篇文章被收在1890年出版的一本康托尔关于超限数的论文选集

中。弗雷格在应约评价这本书时，特意强调了刚刚引用过的这段话。在他收到伯特兰·罗素那封毁灭性的信之前10年的一段引人注目的话中（本章开头已经部分引用过），弗雷格写道：

> 的确如此！这就是将会使这个学说毁于一旦的暗礁。因为无限终将不会从算术中被排挤出去；然而另一方面，它是无法与这种认识论倾向共存的。于是我们可以预见，这个问题将会为一场关系重大的、决定性的斗争搭好舞台。[20]

1918年1月6日，格奥尔格·康托尔因心脏病突发而去世，其时，第一次世界大战的战火仍在蔓延。今天，尽管弗雷格在其隐喻中所预言的斗争很是令人惊奇，但它却并未产生任何决定性的结果。也许这场斗争的最令人称奇的副产品就是阿兰·图灵关于一种通用计算机的数学模型。

第5章　希尔伯特的营救

乔治二世是莱布尼茨的最后一任资助人英王乔治一世的儿子，1737年，他在德国中部莱内河畔的中世纪小城哥廷根建立了一所大学。城市的围墙、几所哥特式教堂以及古老街道上的半木质的房子仍然保存至今。哥廷根大学那令人骄傲的数学传统可以追溯到19世纪，那里产生了像卡尔·弗里德里希·高斯、伯恩哈德·黎曼、勒热纳·狄里克莱和菲利克斯·克莱因这样伟大的数学人物。但哥廷根真正的数学辉煌还是在20世纪。那时，主要是受到大卫·希尔伯特的感召，世界各地的学生都来到这个无可争议的世界数学中心，直到1933年纳粹统治德国所导致的大批流亡为止。

20世纪40年代末，当我本人在读研究生期间，关于哥廷根在20年代的逸事仍然在被一代代的人传诵着。我们听说了卡尔·路德维希·西格尔在容易上当受骗的贝塞尔—哈根身上实施的无休无止的恶作剧。我个人最喜欢的故事是，人们看到希尔伯特日复一日地穿着破裤子，这对许多人而言都是令人尴尬的。把这种情况得体地告诉希尔伯特的任务落在了他的助手理查德·库朗身上。库朗知道希尔伯特喜欢一边谈论数学，一边在乡间漫步，于是便邀请他一同散步。库朗设法使他们两人走过一片多刺的灌木丛，这时库朗告诉希尔伯特他的裤子被灌木丛刮破了。“哦，不是，”希尔伯特回答说，“它几星期前就是这样了，不过还没有人注意到。”正是在20世纪的20年代，希尔伯特掀起了一场令人瞩目的运动，那就是用数学来证明数学本身的合理性。从希尔伯特的运动到阿兰·图灵对计算本性的洞察，其间发生了一连串怪事。

大卫·希尔伯特出生于小城柯尼斯堡的一个新教家庭。柯尼斯堡位于普鲁士东部，以哲学家伊曼努尔·康德的故乡而闻名。1870年，俾斯麦与拿破仑三世的法国交战，随即借着德国的巨大胜利把德国统一成了一个以普鲁士国王为皇帝的帝国，当时希尔伯特才8岁。到了他进入柯尼斯堡大学学习数学时，他在这门学科上的非凡才华已经得到公认，他把数学融入谈话的特有风格也建立了起来。他可以同他的朋友赫尔曼·闵可夫斯基和阿道夫·胡尔维茨一起长时间地漫步讨论数学。[1]

从莱布尼茨和牛顿创立微积分到大卫·希尔伯特成为数学家，在这2个世纪里，不少人都发现极限过程可以漂亮地应用于许多方面。在这些结果中，有很多是通过对符号的纯形式操作来获得的，人们对它们背后的含义并未做过多考虑。但是到了19世纪中期，清算之日还是来临了，需要对符号进行概念理解的问题层出不穷。冲在解决这些问题最前线的则是格奥尔格·康托尔、他的老师卡尔·外尔施特拉斯以及他的朋友理查德·戴德金。

1888年，希尔伯特游历了德国各个主要的数学中心，以同他这个领域的主要人物进行接触。在柏林，他拜访了康托尔的强硬对手利奥波德·克罗内克，他们2年前就已相识。克罗内克是一个大数学家，他的一些著作对于希尔伯特的成就起着相当重要的作用。但是，正如希尔伯特昔日的学生赫尔曼·外尔在半个世纪后的一篇悼文中所写的，希尔伯特认为克罗内克正在用“他的力量和权威迫使数学服从专断的哲学原理”。这些原理使得克罗内克对他那个时代的大部分数学持一种极度否定的态度。克罗内克不仅反对康托尔的超限数，而且也反对外尔施特拉斯、康托尔、戴德金为微积分的极限过程提供一个坚实的基础所做出的全部努力。克罗内克把这些努力斥之为毫无价值。他特别主张对存在性的数学证明应当是构造性的。也就是说，要想让克罗内克接受一个对满足某种条件的、实际存在着的数学对象的证明，它就必须能够提供一种方法来明确地呈现出这个对象。希尔伯特不久将会在著作中挑战这种说法。许多年以后，他将这样向学生来解释这种区分，即在礼堂中所有听讲的学生中（他们中没有人是完全秃顶的），总有一个学生是头发最少的，尽管他没有明确的办法来识别出这个学生。[2]



希尔伯特早期的胜利

这个世界是不断变化的，但有些东西却是不变的。数学家们通常关心的正是找到那些能够在其他事物变化时保持不变的东西。这时，他们谈论在某些变换下保持不变的那些东西。乔治·布尔在一篇早期的论文中开创了对所谓代数不变量的研究。[3] 到了19世纪的最后25年，代数不变量已经成了数学研究的主要焦点之一。人们在代数操作上展开了英雄式的较量，以期解决如何发现不变量的问题。在这种努力中，德国数学家保罗·果尔丹是一位真正的行家，他被其同时代人戏称为“不变量之王”。在研究代数的过程中，果尔丹猜出了一个关于代数不变量结构的简化定理。根据果尔丹的猜想，在考虑一个特定的代数表达式的所有不变量时，总是会有几个主要的不变量，借助它们，所有其他的不变量都可以用一个简单的公式来表达。然而，他的这次努力只是证明了，这一猜想仅仅在一个非常特殊的情况下才是正确的。果尔丹的猜想被当时的数学家视为所面临的主要问题之一，一般认为，只有在操作代数上具有果尔丹那样能力的人才能证明它。在这种情况下，大卫·希尔伯特对果尔丹猜想的证明引发了强烈的震撼。希尔伯特依靠的不是复杂的形式操作，而是抽象思维的力量。

只是在与果尔丹本人会面之后，希尔伯特才发现自己被果尔丹所提出的问题迷住了。他的解决方案花了6个月时间才完成，它建立在一个极具一般性的结论之上，即今天所谓的希尔伯特基本定理，其证明非常直接。利用这个基本定理，希尔伯特证明了，如果假定果尔丹的猜想是错误的，那么就会导致矛盾。这种对果尔丹猜想的漂亮证明不会使克罗内克感到满意，因为它不是构造性的。希尔伯特没有列出其存在性已经确定无疑的一系列主要不变量，这一证明只是表明，如果假定它们不存在将会导致矛盾。然而，运用抽象思维的力量，希尔伯特的证明为即将来临的那个世纪的数学开辟了道路。希尔伯特的证明所揭示出来的更为普遍的观点足以摧毁代数不变量的古典理论。今天，果尔丹主要因其对希尔伯特证明的反应而被人记住。“这不是数学，”他惊呼，“这是神学。”

希尔伯特对果尔丹问题的解决所产生的轰动效应立刻使他成为当时

的第一流数学家。此后，他并没有满足于所取得的成就。在永远离开不变量理论之前，他又对一些细节进行了归整，特别是又给出了一个完全构造性的对果尔丹猜想的证明。 [4] 此外，他还就各种数学主题发表了许多论文。值得注意的是，有一篇小论文无疑带有一种必定为克罗内克所不齿的康托尔的气息。然而，尽管如此多产，希尔伯特实际的职业却并不尽如人意。多年来，他一直在柯尼斯堡大学任编外讲师，依靠他的演讲所获得的微薄收入过活。有一段时间，他只给一个来自巴尔的摩的学生上了整整一门课。在给他的好朋友闵可夫斯基的一封信中，希尔伯特讽刺地说，这里有11个编外讲师竞争同样数目的学生。

1892年，年轻的希尔伯特的生活发生了重大改变。这始于68岁的克罗内克于年底去世，卡尔·外尔施特拉斯将于次年退休。德国数学的封闭的学术生活开始复苏了，这引出了德国数学界的抢座位游戏。在做了6年编外讲师之后，希尔伯特终于在柯尼斯堡大学获得了一个正式的学术职位。也是在同一年，他与自己最心爱的舞伴凯特·耶罗士结了婚。1年之后，他们的儿子弗朗茨出生了。与此同时，哥廷根大学的数学领袖菲利克斯·克莱因决定动员希尔伯特到哥廷根去。1895年春，克莱因的动员成功了，希尔伯特搬到了哥廷根，他将在那里一直待到48年后去世。

如果说希尔伯特对果尔丹问题的出色证明宣告了代数不变量古典理论的终结，那么他应德国数学会之邀所写的内容丰富的《数论报告》（*Zahlbericht* ）则预示着一大批数学成果的诞生。学会本以为这是一篇关于一门相对较新的数学分支——代数数论当前状况的报告，许多数学家都对这个主题感到困惑， [5] 但他们最后得到的却是一篇经过深思熟虑的、从第一原理出发对这一领域进行重新组织的报告。直到半个世纪后我做研究生时，我们仍在饶有兴味地研究它，并且受益匪浅。

当希尔伯特来到哥廷根时，他已经为一大批数学课程做好了上课准备，他在柯尼斯堡任编外讲师时就一直在上这些课。奥托·布鲁曼塔尔是他所指导的69名获得博士学位的学生中的第一个，他在40年后的回忆录中仍然能够清楚地回忆起自己刚到哥廷根时希尔伯特留给他的印象：“这位中等身材的思路敏捷的人留着红色的胡须，穿着非常普通的衣服，看起来一点也不像一个专家……[与其他教授相比]”布鲁曼塔尔是这样描述希尔伯特讲课的：

他讲课紧扣主题，但风格却相当沉闷，不时会重复一些重要的命题。然而，谈话中所包含的丰富内容以及它的清晰却使人忘却了形式。他会介绍一些新的以及他本人曾经做过的工作，但不会特别强调它们。为了使每个人都能理解，他显然动了不少脑筋；他是给学生讲课的，而不是给他自己。[6]

1898年的秋季学期，学生们很吃惊地发现，希尔伯特准备开设一门“欧几里得几何原理”课程。他们原以为他完全沉浸在代数数论中，而从未想到他会对几何学科目感兴趣。这个题目显得很奇怪，因为欧几里得几何学毕竟是一门中学课程。真正的惊讶出现在课程开始之后，这时学生们发现他们听到的是对几何基础的一种全新的发展。这是希尔伯特对数学基础保有深刻兴趣的第一个迹象。我们主要关注的正是这种兴趣。在讲演中，希尔伯特提出了一套几何学公理，以弥补欧几里得的古典处理中的几个漏洞。他强调了这门学科的抽象本性：必须能够说明，那些定理通过纯逻辑就可以从公理中推导出来，而不必受到我们从图形中所“看到”的东西的影响。有一则著名的逸事是这样的：据说他曾经讲，我们可以不说点、线、面，而说“桌子、椅子和啤酒杯”，只要它们遵守公理，定理就必须仍然能够成立。最后，希尔伯特证明了他的公理是一致的，也就是说从公理中导不出矛盾，这样就使他的成就更为牢靠。这一论证表明，他的几何公理系统中的任何不一致都会导致算术中的不一致。因此，希尔伯特所做的工作就是把欧几里得几何学的一致性归结为算术的一致性，而算术的一致性问题则留待它日解决！

面向一个新世纪

1900年8月，在巴黎出席一次国际会议的数学家们都想知道，新的世纪会给这门学科带来什么。此时，令人眩目的成就已经使38岁的大卫·希尔伯特到达了事业的巅峰。在这次会议上，希尔伯特向代表们做了一个特邀报告。在报告中，他列举了23个用当时的方法似乎很难解决的问题，作为对20世纪数学家的一个挑战。[7] 希尔伯特以其特有的乐观主义宣称，所有数学家都深信，“每一个明确的数学问题都必定可以完全得到解决……这一信念……是对我们工作的巨大鼓舞。我们在心中听到了这样的永恒召唤：这里有一个问题，去找出它的答案，你能够通过纯理性找到它。”希尔伯特列出的第一个问题就是康托尔的连续统假设（即没有集合的基数介于自然数集的基数与由所有自然数集所组成的集

合的基数之间）是否为真。在悖论的威胁对克罗内克的否定态度极为有利的这个时期，这是对康托尔的超限数的明确肯定。

第二个问题是希尔伯特对欧几里得几何学公理一致性的证明所留下的问题——为实数的算术建立公理的一致性。以前对一致性的证明都是关于相对一致性的证明，也就是说它们都是把某一个公理集合的一致性归结为另一个集合的一致性。但希尔伯特意识到，对于算术，他已经抵达了逻辑的根底，此时还需要有新的直接方法。这个问题也为希尔伯特创造了一个机会来解释数学上的存在性的含义。尽管克罗内克已经声称，要想证明数学对象存在，就必须提出一种能够构造或展示相关对象的方法，但在希尔伯特看来，存在性只要求能够证明，假设这些对象存在并不会导致矛盾："如果可以通过有限次的逻辑过程证明，赋予一个概念的性质永远不会导致矛盾，那么我们就可以说，这个概念的数学存在性……就被证明了。"根据希尔伯特的说法，既然假定一切康托尔的超限基数所组成的集合的存在会导致矛盾，那么这就表明这样一个集合并不存在。特别是在伯特兰·罗素1902年写给弗雷格的那封毁灭性的信中所说的悖论广为人知以后，人们渐渐地认为数学基础所面临的困难导致了一场危机，算术的一致性问题一直令人头痛。直到20世纪20年代，希尔伯特和他的学生们集中精力研究了这个问题，其后果他们几乎不可能预见到。

希尔伯特在1900年所列出的那些问题使一代又一代的数学家殚精竭虑。它们涵盖了纯粹数学和应用数学中的大量主题，预示了希尔伯特未来将会做出的贡献的广度。在一篇悼念希尔伯特的文章中，赫尔曼·外尔评论说，任何一个解决了希尔伯特所列出的问题之一的人都会成为"数学共同体这一光荣集体中的一员"。1974年，美国数学会主办了一个特别的研讨会（我有幸参加了这个会议），专家们在会上讨论了自那时以来源自这些问题的数学进展。会后出版的论文集足足有600多页，希尔伯特问题的丰富内涵由此可见一斑。[8]

克罗内克的幽灵

随着伯特兰·罗素公开了他在看似浅易的推理中发现的那个悖论，许多数学家对康托尔的超限数以及基础研究的整个方向产生了前所未有的疑虑。正如我们所看到的，当弗雷格收到包含罗素悖论的那封信后，

他径直放弃了自己毕生的工作。不知此时他是否还记得自己10年前曾经做过的预言：

> 由于无限在算术中的地位终究是不容否认的……于是我们可以预见，这个问题将会为一场关系重大的、决定性的斗争搭好舞台。[9]

虽然弗雷格和康托尔的朋友戴德金都从斗争中退了出来，但介入这场斗争的勇士却不乏其人。在20世纪的早期，两位世所公认的最伟大的数学家希尔伯特和昂利·庞加莱都参与了这场争论，但他们所持的观点却并不相同。1900年以后，接下来的一届国际数学家大会于1904年召开，此时距罗素公布他的悖论已经有2年之久。希尔伯特在会上所做的报告中略述了算术的一致性证明可能采取的形式，并且明确提出了自己解决危机的方法。[10] 他没有忘记指出，这个证明可以被拓展到把康托尔的超限数也包括进来。庞加莱不久就发现，希尔伯特犯了循环论证的错误：这个证明所要辩护的方法被用在了对这些方法不可能导致矛盾的证明中。又过了一些年，希尔伯特才有能力回应这个反对意见。庞加莱发现他所谓的“康托尔主义”并非毫无用处，但他坚持说，“（给定的、完成的）实无限并不存在。康托尔主义者忘记了这一点，他们陷入了矛盾。”[11] 这里，庞加莱与高斯在80多年前所写的话（上一章中已经提到）相呼应：“我极力反对把无限当成一种完成的东西来使用，这在数学上是绝不能允许的。”康托尔毕生的伟大工作都是对这一传统的英勇无畏的挑战。

伯特兰·罗素并没有从战场中退却。他一直希望能够发展出一种符号逻辑体系，利用这种体系，就可以实现弗雷格把算术还原为纯逻辑的计划而不会导致悖论。在向其同时代人解释他的工作的过程中，他极大地得益于意大利逻辑学家朱塞佩·皮亚诺所创立的符号系统（特别是第3章中所介绍的那个系统），这个系统远比弗雷格的系统易于理解。庞加莱在抨击罗素的工作时尖刻地指出：

> 很难理解，当我们写下⊃的时候，“如果”一词便获得了当我们写“如果”时它并不具有的一种优点。[12]

庞加莱也没有忘记指出，严肃地看待罗素的工作将有可能把数学还原为纯粹计算。（莱布尼茨之梦！）

在嘲笑这一观念时，他说：

很容易理解，要证明一个定理，知道它是什么意思并

不是必需的，甚至也不是有用的……我们可以想象这样一台机器，我们从一端输入公理，从另一端就可输出定理，就像芝加哥那台传说中的机器一样，猪活着进去，出来的就是火腿和香肠。数学家和这些机器都不需要知道自己正在做什么。[13]

伯特兰·罗素挽救弗雷格计划的努力体现于罗素与阿尔弗雷德·诺斯·怀特海合写的三卷本巨著《数学原理》（*Principia Mathematica* ）中（出版于1910—1913年）。这部著作从弗雷格《概念文字》的纯逻辑开始，以清楚明白的数学为结束，中间是简单而直接的步骤，它完全体现了庞加莱的芝加哥机器的精神。悖论通过一种精心设计的、使用起来很不方便的分层结构而得以避免。事实上，在这种结构中，任何一个集合都只能拥有来自同一层的成员。这种分层大大削弱了普通数学的能力，以致一条特别可疑的还原公理被提了出来，以穿透层与层之间竖立起来的壁垒。[14] 此外，《数学原理》还因一种背后的混淆而受到损害。虽然弗雷格已经清楚地认识到他正在处理两种层次的语言——他正在构造的一种新的形式语言以及可以谈论这种新语言的日常语言，但怀特海与罗素的这部著作在这个问题上却不够清楚，它把这两个层次混在了一起。[15] 这就意味着，在希尔伯特看来如此重要的整个结构的一致性问题在罗素的语境下甚至就不会出现。尽管如此，《数学原理》仍然是一个里程碑式的成就，它一劳永逸地证明了，在一个符号逻辑系统中对数学进行完全的形式化是绝对可行的。

当伯特兰·罗素正在努力为古典数学的广度寻找一种逻辑基础以避免悖论时，一位优秀的年轻荷兰数学家L.E.J.布劳威尔确信，这一切几乎都是错误的，都应被抛弃。布劳威尔1907年所写的博士论文表现出对康托尔的超限数以及当时大部分数学工作的强烈敌意，以至于人们认为他骨子里全都是克罗内克的精神。1905年，布劳威尔从数学研究中抽空出版了一本小书《生活、艺术和神秘主义》（*Life, Art and Mysticism* ），书中充斥着浪漫的悲观主义情绪。在描绘了这个“忧伤的世界”中的生活之后，这个忧郁的年轻人总结说：

放眼望去，这个世界满是不幸的人，他们想象自己能够拥有财

产……同时滋生一种对知识、权力、健康、荣耀和愉悦的不知餍足的欲望。

只有那些认识到自己一无所有、无法拥有任何东西、安全是不可企及的人，那些完全隐退、牺牲一切的人，那些不知道任何东西、不渴望任何东西也不想知道任何东西的人，那些放弃和忽视一切的人，才能得到一切：自由的世界向他开启了，这是一个没有痛苦的沉思的世界——一个一无所有的世界。[16]

尽管布劳威尔对自我弃绝的生活大唱赞歌，但他却发动了一场自以为是的、从头开始重建数学的斗争，以使他的哲学信念得到满足。他本可以轻易地选择一个传统的数学主题，但他还是决定写一篇关于数学基础的博士论文。[17] 他的导师很不情愿地答应了，但由于他的这位高明的学生坚持要把那种奇怪的不相干的观念带入其博士论文，这位导师惊骇地写道：

至于我是否可以原封不动地接受第2章，我又一次做了考虑。但说老实话，布劳威尔，我不能接受。我发现它充斥着对生活的悲观主义和某种神秘主义态度，它既不是数学，与数学基础也没有任何干系。[18]

在布劳威尔看来，数学就存在于数学家的意识中，它最终导源于时间这个“数学的原初直观”。真正的数学在数学家的直观中，而不在语言的表达中。数学非但不是逻辑（如弗雷格和罗素所主张的），逻辑本身倒是来源于数学。在布劳威尔看来，康托尔相信自己已经找到了不同大小的无限，这纯粹是无稽之谈，他的连续统问题也是无足轻重的；希尔伯特宣称一致性就是数学上的存在性所需要的一切，这是错误的。事实恰恰相反：

在数学中存在就意味着：由直观构造出来；某种特定的语言是否一致，这个问题不仅本身并不重要，而且也不是对数学上的存在性的一种检验。[19]

克罗内克认为，确立数学上的存在性的唯一有效方法就是构造。布劳威尔比这走得更远，他坚决反对把亚里士多德的排中律这条基本的逻辑定律（它断言任何命题或者为真，或者为假）应用于无限集。[20] 在

布劳威尔看来，有些命题既不能说为真，也不能说为假；对于这些命题，已知的方法还不能判定这一点。希尔伯特在对果尔丹问题的最初证明中使用了排中律，这是数学家的通常做法：他证明了否定这个猜想将会导致矛盾。在布劳威尔看来，这样一个证明是不可接受的。

完成了博士论文之后，布劳威尔决定暂时把他备受争议的观点搁置起来，把注意力集中到显示他的数学才能上。他选择的竞技场是刚刚兴起的拓扑学领域。他得到了几项深刻的结果，其中包括那条重要的不动点定理。[22] 1910年，当29岁的布劳威尔发表这个基本定理时，他已经赢得了希尔伯特的称赞。大卫·希尔伯特对此印象很深，他甚至还邀请这个年轻人加入他那份著名的杂志《数学年鉴》（*Mathematische Annalen*）的编委会，后来他对这次邀请后悔不已。在1912年获得了阿姆斯特丹大学的一个常任教席之后（这得到了希尔伯特的举荐），布劳威尔觉得可以回到他那革命性的设想了，他现在把它称为直觉主义。

赫尔曼·外尔是希尔伯特的得意门生，他是那个世纪最伟大的数学家之一，也是后来接任希尔伯特在哥廷根职位的人。他的兴趣涵盖了数学、物理学、哲学甚至艺术。但是令希尔伯特非常沮丧的是，外尔确信，外尔施特拉斯、康托尔和戴德金已经建立的处理极限过程的基础是摇摇欲坠的，他无法接受所有这些所基于的实数系统。他在一段著名的话中宣称，整个大厦“都建立于沙堆之上”。[21] 外尔重建实数连续统的努力最终并没有令他本人满意，当他得知布劳威尔对此的看法时，不禁为之着迷。“……布劳威尔，这就是革命。”他宣称。在希尔伯特看来，这是太过分了，他也许想到了那句“还有你吗，布鲁图斯？”（Et tu, Brute？）[23] 19世纪20年代，德国确实是大变革时期。德国在第一次世界大战中已经战败，被迫签订了《凡尔赛条约》。在德皇退位之后上台的社会民主党政府已经被严重的经济问题以及左翼和右翼让它下台的呼声搞得焦头烂额，各个方面都传来激烈的言辞。在这种使人头脑发热的气氛中，希尔伯特在1922年所做的一篇讲演中把他过去学生的离弃比作背叛：

> 外尔和布劳威尔的所作所为归根结底是在步克罗内克的后尘！他们要将一切他们感到麻烦的东西扫地出门，以此来提供数学基础，并且以克罗内克的方式宣布禁令。但这将意味着肢解和破坏我们的科学，如果听从他们所建议的这种改革，我们就有可能丧失大部分最宝贵的财富。外尔和布劳威尔把无理数的一般概念、函数甚

至是数论函数、康托尔的超限[基]数等等都宣布为不合法。无限多个自然数中总有一个最小的数这一定理，甚至是逻辑上的排中律，比如断言或者有有限多个素数，或者有无限多个素数：这些都成了明令禁止的定理和推理模式。我相信，正如克罗内克不能废除无理数一样……外尔和布劳威尔今天也不可能获得成功。不！布劳威尔的[纲领]并不像外尔所相信的那样是在进行什么革命，他只不过是在重演一场有人尝试过的暴动，这场暴动在当初曾以更凶猛的形式进行，结果却彻底失败了。何况今日，由于弗雷格、戴德金和康托尔的工作，数学王国已经是武装齐备，空前强固。因此，这些努力从一开始就注定要遭到同样的厄运。[22]

像无数的欧洲人一样如果我们注意到希尔伯特言辞中浓烈的火药味，很可能以为他在狂热地欢呼即将来临的1914年的战争呢。但事实并非如此。从一开始，他就公开表示自己认为战争是愚蠢的。1914年8月，93位著名的德国知识界人士联名向“文明世界”发表了一份宣言，以回应英国、法国和美国对德军入侵比利时的愤慨。这份宣言称：“说我们已经非法地侵犯了比利时的中立……我们的军队已经残忍地把卢汶夷为平地，这些都不是真的。”希尔伯特被要求签字，但他拒绝了，他说他不知道这些指责是否是真的。1917年，即希尔伯特公开指责外尔和布劳威尔的5年前，当血腥的堑壕战仍在吞噬一代欧洲人的时候，希尔伯特发表了一篇给刚刚去世的法国大数学家伽斯东·达布写的悼文。这篇文章刊出后，一群凶悍的学生聚集到他的住宅前，要求他收回这篇悼念“敌人数学家”的文章。希尔伯特拒绝了，他跑到校长办公室威胁说要辞职，除非就这些学生的无理行为向他做官方的道歉。他很快就收到了这样的道歉。[23] 当优秀的年轻数学家埃米·诺特被任命为哥廷根的编外讲师时，有人认为这将导致一个女人成为教授和大学评议会成员，这时希尔伯特说：“我看不出一个候选人的性别为什么会成为不能让她当讲师的理由，大学评议会毕竟不是澡堂。”[24] 1917年9月，当德国与它的邻国法国正竭尽全力杀戮对方的公民时，希尔伯特在苏黎世发表了一篇题为“公理化思想”的讲演，开头是一句富有挑衅意味的话：

正像在民族的生活中那样，一个民族只有在和它的所有邻邦都处理好关系时才可能繁荣，国家的利益不仅要求每一个民族都服从命令，而且也要求民族之间的关系能够被妥善处理，在科学的生活中也是这样。[25]

元数学

算术的一致性问题在希尔伯特1900年在国际数学家大会上所做的报告中排在第二位。然而只是到了20世纪的20年代，希尔伯特才严肃地提出了自己处理这个问题的方法。他的学生威廉·阿克曼以及助手保罗·贝尔奈斯都与他通力合作，约翰·冯·诺依曼也做出了贡献。 [24] 希尔伯特从怀特海—罗素《数学原理》的逻辑系统开始，最初也是按照弗雷格—罗素的目标想用纯逻辑的术语来定义数。但他很快就不得不放弃这个目标，不过仍然认为他们所创立的符号逻辑是至关重要的。在希尔伯特的新纲领中，数学与逻辑将通过一种纯形式的符号语言被发展出来。这样一种语言可以从“内部”和“外部”来看。从内部看，它就是数学，每一步演绎都可以完全弄清楚。但是从外部看，它仅仅是许多公式和符号操作，它们可以在不考虑意义的情况下进行处理。这里的任务是要证明，从这种语言中导出的任何两个公式都不会彼此矛盾，或者等价地说（正如后来所表明的情况那样），像1=0或0≠0这样的公式是不可能导出的。

庞加莱和布劳威尔所提出的批评必须要面对这样的挑战：如果一种一致性证明想要确保的方法恰恰就是它所依赖的方法，那么从这样一种证明中是不可能产生任何有价值的东西的。希尔伯特的大胆想法是一种全新的数学，他称之为元数学或证明论。一致性证明将在元数学内部完成。尽管在形式系统内部，每一种数学方法都可以被不加限制地运用，但元数学方法却要严格限于那些被希尔伯特称为“有限性”（finitary）的无可争议的方法。这样希尔伯特就可以嘲弄布劳威尔和外尔说：我已经证明，数学家们使用通常的那些方法是不会导致矛盾的，我证明这个结论所使用的方法甚至就是你们所赞同的那些方法。或如冯·诺依曼所指出的：“可以这样说，证明论在直觉主义的基础上把古典数学建立起来，并以这种方法把直觉主义归于荒谬。”[26]

在他的方法所要营救的数学“宝藏”中，希尔伯特强调指出要把康托尔的超限数包括进来。关于这一点他说：“在我看来，这是数学领地所开出的最令人惊叹的花朵，它是人类纯理性活动的最高成就之一。”[27] 他拒不接受布劳威尔和外尔的批评，并且宣称：“没有人能把我们从康托尔为我们建造的乐园中赶走。”[28] 虽然布劳威尔已经准备承认希尔伯特纲领很可能会取得成功，但他依然初衷不改：“……这样也得不到任何数学价值：对于一个错误的理论来说，即使它不能被任何矛盾所反

驳，它依然是错误的理论，这就如同一种罪恶的政策就是罪恶的一样，即使它不能被任何法庭所禁止。”[29]

当希尔伯特使用半合法的方法把布劳威尔从《数学年鉴》的编委会中撤下时，他们之间的舌战已经上升到了行为，这使阿尔伯特·爱因斯坦开始抱怨“这场蛙鼠之战”。[30] 希尔伯特及其合作者与布劳威尔和外尔之间的争论当然都植根于关于知识本性的基本哲学问题。事实上，双方的观点都受到了伊曼努尔·康德的深刻影响。然而，与大多数哲学争论不同，希尔伯特和布劳威尔所持的立场是以纲领的方式表现出来的，这导致了非常具体的问题，于是也就潜藏着被后来的事件所推翻的可能。

布劳威尔所面临的主要问题是如何真正实现他的纲领中所说的对数学的重建，如何使数学家们相信，他们可以不用古典的实数连续统，不用排中律，也仍然不会丧失他们那些最宝贵的财富。然而，布劳威尔所提出的直觉主义数学却面临着外尔后来所说的“一种几乎令人无法承受的尴尬……”，它并没有使多少人皈依。[31] 尽管布劳威尔从未放弃过自己的观点，但他愈来愈感孤立，其晚年是在“莫须有的经济顾虑和对破产、迫害和疾病的妄想的恐惧”中度过的。1966年，当他85岁时，他在通过他家门前那条街道时被一辆车撞死。[32] 也许这个故事中最大的讽刺是，直觉主义之所以能够存活下来，并不是因为它像布劳威尔所设想的那样，成了数学工作者的正确实践，而是因为对包含他的思想的形式逻辑系统的研究。[33] 其中有些系统实际上已经成了实现形式演绎的计算机程序的基础。[34]

当然，希尔伯特纲领所提出的主要问题就是它最初的问题：算术的一致性问题。阿克曼和冯·诺依曼研究了这个问题，并且取得了部分成果。当时认为这不过是一个打磨技巧以得到完整结果的问题。1928年，希尔伯特和他的学生阿克曼出版了一本薄薄的逻辑课本，这本书是基于希尔伯特（在贝尔奈斯的协助下）自1917年以来的授课内容写成的。书中提出了两个关于弗雷格《概念文字》的基本逻辑（后来被称为一阶逻辑）的问题。从某种意义上说，这两个问题已经流传了一段时间，但正是希尔伯特对逻辑系统可以从外部进行研究的这种眼光，才造就了表达它们的清晰形式。其中一个问题是证明一阶逻辑的完备性，即任何一个从外部看来有效的公式都可以只用课本中提出的规则从系统内部导出。第二个问题以希尔伯特的“判定问题”（*Entscheidungsproblem*）而闻

名，即对于一个一阶逻辑的公式，如何找到一种方法，可以在定义明确的有限的步骤内判定这个公式是否是有效的。我们在第7章中将会看到，作为需要数学家解决的具体问题，这两个问题的解决在20世纪出现了希望，而在17世纪还只能是莱布尼茨的梦想。

在1928年，希尔伯特在波伦亚举行的一届国际数学家大会上做了讲演。除非受到国际形势的影响，这些会议每4年举行一次。当然，1916年没有举行会议。1920年和1924年的会议都举行了。但是战后的痛楚是如此强烈，以至于德国人没有受到邀请。是希尔伯特坚持要德国数学家接受参加1928年国际数学家大会的邀请，而不是像路德维希·比贝尔巴赫（后来成为纳粹分子）和布劳威尔那样，希望联合抵制会议作为对《凡尔赛条约》的抗议。在讲演中，希尔伯特提出了一个关于形式系统的问题，这个系统建立在把一阶逻辑规则应用于现在被称为皮亚诺算术或PA（以意大利逻辑学家朱塞佩·皮亚诺的名字命名）的自然数公理系统的基础之上。希尔伯特希望能够证明PA是完备的，也就是说，任何一个可以在PA中表出的命题，或者可以在PA中被证明为真，或者可以在PA中被证明为假。2年后，这个问题被一个名叫库尔特·哥德尔的年轻逻辑学家解决了，但答案完全不像希尔伯特所预料的那样。事实上，后来的情况表明它对于希尔伯特纲领有着巨大的破坏作用。

灾难

希尔伯特的传记作家把他的妻子凯特描绘成一个聪慧而有判断力的人、她丈夫的贤内助（他的许多手稿都是她抄写的）、一位母亲以及向年轻的数学家们传授生活智慧的人，对于那些年轻的数学家，希尔伯特的房间大门永远是敞开着的。希尔伯特自认为是一个尘世中人，他曾经嘲弄说，自己最好的休假是和一位同事的妻子度过的。他从未厌倦逢场作戏，只要情况允许，他就会试图与最漂亮的年轻女性跳舞。他的“激情”是如此声名狼藉，以至于在一个快乐的生日宴会上，有人即兴作诗，讲他的“爱情”可以用字母表中的每一个字母代表不同的人。但是到了字母“K”时，每个人都被难住了。这时凯特说：“你至少可以想起我一次了。”于是就有了下面的几句诗：

感谢上帝，

她将不会争吵。

“谁在乎啊，”凯特道，

她是他的妻子。

对这对夫妻来说，他们的儿子弗朗茨一直是痛苦的来源（在许多方面）。他虽然在身体上与父亲极为相似，但在精神领域却没有任何相似之处。尽管他竭力掩饰自己，但事实很清楚，弗朗茨是一个心理极不正常的人，以至于最后不得不被送到医院。父亲对这个悲剧的反应是他不再有一个儿子了，而母亲却不这么想。

1929年，一幢新房子建成了，这里将成为哥廷根数学研究所的所在地。资金由洛克菲勒基金会和德国政府提供，这在很大程度上要归功于理查德·库朗卓有成效的外交努力。但是哥廷根成为世界数学研究中心的日子很快就要结束了。当希尔伯特1930年退休时，赫尔曼·外尔被邀请接任他的职位。同年，希尔伯特被他的出生地柯尼斯堡授予“荣誉市民”称号。那年秋天，他应邀于柯尼斯堡在德国科学家和医生协会的会议上做了一场特殊的讲演。希尔伯特特地选了一个较具一般性的主题：自然科学与逻辑。在这篇内容广泛的演说中，他强调了数学在科学中以及逻辑在数学中所扮演的极端重要的角色。带着他那一贯的乐观主义，他坚持说不存在不能解决的问题。他以下面的话作为结尾：[35]

Wir müssen wissen（我们必须知道），

Wir werden wissen（我们将会知道）。

就在希尔伯特发表演讲前几天，一个关于数学基础的研讨会在柯尼斯堡召开了。演讲者有布劳威尔的学生和追随者A.海丁、哲学家鲁道夫·卡尔纳普以及（代表希尔伯特证明论纲领的）约翰·冯·诺依曼。在会议结束时的圆桌讨论会中，一个名叫库尔特·哥德尔的羞怯的年轻人（我们下一章的主题）不动声色地做了一项声明，对于那些领会了其要旨的人们来说，这标志了基础研究中的一个新的时代。冯·诺依曼立刻意识到一切都结束了——希尔伯特的纲领是不可能成功的。当希尔伯特得知哥德尔的声明时，他最初多少有些生气，因为在他看来，这是对他的“我们将会知道”的当头一棒。但是当贝尔奈斯用1934年和1939年出版的两大卷著作写下了希尔伯特证明论的成就时，哥德尔的工作起了重要

作用。[36]

1932年，希尔伯特的70岁生日如期在新落成的数学研究所举行了庆祝活动。席上有烤面包和音乐，当然还有舞蹈，这位老人对跳舞仍然乐此不疲。也是在1932年，沮丧之情达到了顶点，纳粹在国会选举中大获全胜。在接下来的1月份，希特勒被任命为总理，德国科学的崩溃接踵而至。犹太人被禁止授课，他们纷纷地逃往了国外。尽管理查德·库朗在一战期间参加了德军的战斗，但他还是被那个他已经付出了如此多心血的数学研究所拒之门外。最终他在纽约大学创建了另一个数学研究所，这个以他的名字命名的研究所坐落在纽约市的格林威治村。尽管赫尔曼·外尔是一个“雅利安人”，但他仍然觉得德国的形势无法忍受，于是就接受了普林斯顿高等研究院的一个职位，加入了阿尔伯特·爱因斯坦的队伍。[25]

希尔伯特似乎被这种新的政治形势弄糊涂了——他一方面公开抨击政治制度，即使是在这样做越来越危险的时候，另一方面又不能理解自诩的德国法律体系为什么无法抵御任何袭击。在一次聚会中，希尔伯特问他的第一个博士生布鲁门塔尔在教什么课。当得到的回答是他不再被允许教课时，这位老人义愤填膺地问他为什么不去诉诸法律。布鲁门塔尔本人去了荷兰，但是当德国人1940年入侵荷兰时，他陷入了绝境。1940年，他死在了今天位于捷克共和国的特里西恩施塔特（Theresienstadt）的臭名昭著的犹太人区。

希尔伯特于1943年去世，此时二战的硝烟还未散去。凯特也于2年之后辞世。希尔伯特的墓碑上刻着这样两行字：

Wir müssen wissen（我们必须知道），

Wir werden wissen（我们将会知道）。

第6章　哥德尔使计划落空

1952年秋的一天，即我携妻子弗吉尼娅到普林斯顿高等研究院作为期2年的访问研究后不久，我们驱车沿一条古老的小路赶往研究院。此时，两个正在漫步的怪人不知不觉走到了我们车前，并挡住了我们的去路。其中高一点的那个男人穿得甚是邋遢，而另一个则着一身很合体的西服，夹着一个公文包。当我小心地驾车经过他们时，我们认出了那是阿尔伯特·爱因斯坦和库尔特·哥德尔。“爱因斯坦和他的律师。”弗吉尼娅笑道。

这对好朋友不仅仅是在衣着上有所不同。1952年的总统大选过后，爱因斯坦曾断言：“哥德尔是完全疯了……他竟然把票投给了艾森豪威尔。”[1] 对于自由主义者的爱因斯坦来说，把票投给共和党人简直是不可想象的。他们两人对许多基本哲学问题的看法也相去甚远。在提出狭义相对论的过程中，爱因斯坦深受恩斯特·马赫怀疑论的实证主义的影响，尤其是马赫对康德先验时空观的批判。康德的学说认为，我们关于时间和空间（尽管是客观的）的观念是独立于经验观察的。哥德尔在青年时就开始阅读康德的著作，而且终生都保持对德国古典哲学家（尤其是莱布尼茨）的著作的浓厚兴趣。事实上，在一份尚未发表的哥德尔的遗稿中，他认为相对论如果被恰当理解，就可以被看作是证实了康德关于时间本性的某些观点。[2] 与弗雷格和康托尔对实证主义的局限性的不满相呼应，哥德尔承认，正是通过拒斥那些思想，他才有可能看到被其他逻辑学家所忽视的联系，并且做出那些伟大的发现。[3]

在哥德尔1978年去世之后，一个哥德尔协会在维也纳成立了。该协会致力于逻辑以及计算机科学的相关领域中的研究，并且定期在维也纳举行会议。1993年8月，协会在捷克共和国的布尔诺召开会议，纪念哥德尔诞辰87周年。除了既定的科学议程之外，这次会议还特地举行了一个纪念仪式，由布尔诺的市政要人在哥德尔童年时的故居安放纪念牌匾。我仍清楚地记得那时的情景：我们打着伞，站在初秋时节略有些寒意的细雨中；开始是一些捷克语的演讲，随后，一支身着色彩艳丽的民族服装的乐队演奏了几首曲子。

1906年，库尔特·哥德尔出生于布尔诺。那时的布尔诺仍是奥匈帝国的一部分。由于某种原因，伯特兰·罗素一直认为哥德尔是犹太人。事实上，他母亲一家都是新教徒，而他的父亲则在名义上是天主教徒，不过他们平时都很少去教堂。哥德尔一直在德语学校读书，他早年的学业情况被极为完整地保留下来了，这得益于他那一丝不苟的习惯和保存旧物的爱好。从成绩单可知，他在所有的科目中都获得了高分，而他的作业本则显示出当时课程的繁重。8岁那年，哥德尔患了风湿热，这次患病并没有留下什么身体上的后遗症。但此后哥德尔就成了终生的疑病症患者，这很可能就是此次患病的后果。他的哥哥鲁道夫曾说：当库尔特·哥德尔还是个孩子时，就表现出了一些精神上不稳定的迹象。[4]



第一次世界大战后奥匈帝国解体，在新成立的捷克斯洛伐克共和国，操德语的人群变成了少数民族。哥德尔一家发觉这正是他们所处的境况。维也纳坐落在布尔诺南面的108千米处，那里是德语区，有着相当好的大学。于是，鲁道夫和哥德尔很快就决定去那里求学。在以近乎完美的成绩结束了在布尔诺的中学学习之后，哥德尔于1924年秋去了维也纳。在此之前，鲁道夫已经作为一个医科学生到了那里。现在，他们两人可以同居一室了。虽然哥德尔最初打算学习物理学，但在听了一些数论课程之后，整数中所蕴含的那种样式之美使他意识到，数学才是自己真正的追求。

奥地利共和国建立在一战后奥匈帝国的废墟之上，但仅仅过了20年，它就在1938年被纳粹德国吞并了。那是一个混乱喧嚣的年代，这个国家经常游离于内战的边缘。社会民主党的"红色"维也纳与极度保守的乡村之间的冲突一触即发。正是在这种甚嚣尘上的气氛中，著名的维也纳学派生长繁荣起来。怀特海和罗素的工作已经为数学发展出了一种人工语言，它可以将定理的证明表示成纯符号的形式演算。维也纳学派成立于1924年，它由一群哲学家和科学家组成，他们继承了马赫和亥姆霍茨的经验论—实证论传统。我们在第4章中曾经提到，康托尔和弗雷格曾经严厉地抨击过这些思想。维也纳学派对传统的形而上学深恶痛绝，他们深信，哲学的一个主要目的就是发展出类似于怀特海—罗素那样的符号系统，并对其进行研究，这些系统不仅可以包含数学，而且也可以包含经验科学。1936年，该学派的创始人莫里茨·石里克被他以前的一个疯狂的学生枪杀，而纳粹则因莫里茨·石里克可能的左翼立场而宣称该谋杀是正当的。维也纳学派的另一些主要人物还有鲁道夫·卡尔纳普和汉斯·哈恩。鲁道夫·卡尔纳普曾师从弗雷格，汉斯·哈恩则会成为哥德尔的首席教师。[26]

克罗内克的幽灵的回归

当伯特兰·罗素关于数学基础的思想在三大卷《数学原理》中具体成型时，他的学生，那个才华横溢而又富于狂想的维特根斯坦也因那本薄薄的只有75页的《逻辑哲学论》而为世人所知。这两位哲学家的思想在维也纳学派举行的讨论会上扮演着重要的角色。1926年，经汉斯·哈

恩的引荐，哥德尔开始参加这些讨论会，他发觉自己对于所讨论的内容并没有多大兴趣。但即便如此，罗素所阐述的全部的数学都可以用一个形式逻辑系统来表示，以及维特根斯坦所强调的在语言内言说语言的问题，这些肯定影响了年轻的哥德尔的研究方向。维特根斯坦所关注的这些东西与希尔伯特的立场相当一致，他们都认为形式逻辑系统不仅可以在系统内部表达数学推理，而且还可以从系统外部用数学方法加以研究。

在希尔伯特在哥廷根所教授的逻辑课程中，他所采用的逻辑演绎的基本规则源自弗雷格的《概念文字》以及怀特海与罗素合著的《数学原理》。在他1928年的逻辑教科书中（与他的学生威廉·阿克曼合著），希尔伯特提出了在这些规则之间是否存在间隙的问题，也就是说，演绎推理应当是正确的，但规则本身却并不足以保证从前提能够得出结论。他相信并不存在这样的间隙，但他要求对规则本身是完备的进行证明。哥德尔选择了这个问题作博士论文。虽然他很快就得到了希尔伯特所想要的结果，但这其中却不乏反讽的意味。哥德尔所运用的技巧是当时的逻辑学家们相当熟悉的，但为什么他们没有得出这一成果呢？由于布劳威尔—外尔的责难，再加上希尔伯特在元数学的研究中默许了这些责难，这些都影响了他们，束缚住了他们的手脚。我们很快就会看到这一点。

逻辑演绎是从前提到结论的过程。当我们使用弗雷格—罗素—希尔伯特的符号逻辑时，每一个前提和结论皆由一个逻辑公式来表示，也就是相当于一个符号串。[5] 这些符号中有些表示纯粹的逻辑概念，有些只是标点符号，还有一些则是指相关的特定主题。下面是一个逻辑推理的例子：其中前两行是前提，第三行则是结论。

任何一个恋爱中的人都是快乐的。

威廉爱着苏珊。

威廉是快乐的。

运用第3章所介绍的逻辑符号系统，我们可以把它翻译成逻辑语言：

$$(\forall x)\ ((\exists y)\ L\ (x, y) \supset H\ (x))\quad (*)$$

$$L\ (W, S)$$

$$H\ (W)$$

在这个推理中所使用的逻辑符号是⊃，∀和∃，其含义见下表：

⊃如果……，那么……

∀所有

∃某个

字母x， y是变元，它们（就像代词一样）表示被考察人群中的任一个体，而其他符号L， W，H和S所具有的含义则与特定的主题相关。如下所示：

L=恋爱关系

H=快乐这一属性

W=威廉

S=苏珊

由此我们可以把这一推理表示成如下形式：

对于所有的x，如果有一个y为x所爱，则x是快乐的。

威廉爱着苏珊。

威廉是快乐的。

说这个推理是有效的就意味着，无论我们所选择的这些个体所处的群体是什么，无论用字母L表示这些个体间的什么关系，无论用字母H表示这些个体所具有的什么属性，以及无论选择把哪些特定的个体指定给字母W和S，只要我们在推理时保证两个前提皆为真陈述，那么结论就必然为真。为了进一步阐明什么叫作一个推理是有效的，我们不妨用一种完全不同的主题对前面的符号推理做出解释：

食肉动物有着锋利的牙齿。

狼捕食羊。

狼有着锋利的牙齿。

为了说明这个例子也可以用前面的符号推理（*）来表示，我们可以用变元x, y来表示哺乳动物的任意种类，并对其他字母做如下解释：

L=物种之间的捕食关系

H=有锋利的牙齿这一属性

W=狼

S=羊

于是，这一符号推理就可以表示为：

对于所有的x，如果有一个y是被x捕食的，

则x有锋利的牙齿。

狼捕食羊。

狼有锋利的牙齿。

以上解释了推理是有效的是什么含义。希尔伯特希望证明，任何一个有效的推理都可以用弗雷格—罗素—希尔伯特的规则从前提到结论逐步得到证明。换句话说，希尔伯特期望这样一种证明：如果一个推理具有如下属性，

> 不论对公式中的字母做何种解释，只要其前提是真陈述，则它的结论就是真的。

那么我们就可以用弗雷格—罗素—希尔伯特的规则从前提推演出结论。哥德尔在其博士论文中成功地给出了希尔伯特所期望的证明。

哥德尔对自己的证明的解释是直接而又清楚的，这成了他后来所发表的作品的特色。然而，尽管这项重要性随着时间流逝而日趋明显的成果给人以深刻的印象，但他所使用的方法却并没有什么新颖之处，当时的逻辑学家对此都很熟悉。这就使人好奇希尔伯特、阿克曼和贝尔奈斯所组成的强大团队为什么就不能得出这样一个证明。的确，哥德尔在许多年之后评论说，这条定理是挪威逻辑学家特拉尔夫·司寇伦1922年所写的论文的结果的“近乎平凡的推论”。司寇伦的论文比哥德尔的博士论文早了6年（不过哥德尔和哈恩可能对这篇论文都不熟悉）。在一封写于1967年的信中，哥德尔回忆起20世纪20年代，他谈道“当时逻辑学家的盲目……的确很令人奇怪。”不过他又接着说道：

> 我认为这也不难解释。这是因为那个时代的人们普遍缺乏元数

学和非有限性推理所要求的认识论态度。[6]

随着布劳威尔—外尔对非有限性推理提出批评（前面的章节已经讨论过），以及希尔伯特规定在他的元数学中只允许有限性推理，从外部对形式逻辑系统进行研究必须严格限于有限性方法，这至少已经得到了默认，布劳威尔不可能拒绝此种方法。[7] 然而事实上，哥德尔的完备性定理如果不使用非有限性方法就不可能获证。在没有与希尔伯特纲领的目标及其方法论限制相左的情况下，哥德尔解释了为什么非有限性方法在这种情况下可以被允许：

> ……并不是关于数学基础的争论导致了这里所处理的问题（比如数学的一致性问题就是这样）；而是说，即使“幼稚的”数学就其内容而言是正确的从未受到过质疑，这个问题也本可以在这种幼稚的数学中被有意义地提出来（一致性问题与此不同），这就是为什么证明方法上的限制在这里并不像在其他数学问题处那般紧迫。[8]

因此，当哥德尔接受希尔伯特在数学研究中所加的有限性方法的限制以确保数学基础的稳固时，他发现在数理逻辑的工作中没有必要给自己穿上这样一件紧身衣，它并不是这项工作的一部分。

不可判定命题

在希尔伯特于1900年所列的著名问题中，第二个问题是对实数算术的一致性进行证明。当时，没有人知道这样一个证明会是什么样子，特别是不知道它如何才能摆脱循环论证，也就是说，如何在证明中避免用到证明所要证明其为正当的方法。就像我们在前面的章节中所看到的那样，希尔伯特在20世纪20年代介绍了他的元数学纲领：一致性有待证明的公理将被包含在一个形式逻辑系统之内，而证明仅仅是有限数目的符号的一种排列而已。接着，证明这个系统的无矛盾性就要用到希尔伯特所说的有限性方法，该方法甚至比布劳威尔愿意接受的还要严格。当哥德尔完成了他的博士论文转而研究这些问题时，希尔伯特纲领的胜利似乎遥遥在望。

1928年，希尔伯特曾在波伦亚所举行的国际数学家大会上谈到了这

个系统，今天该系统被称为皮亚诺算术（PA），它包含了关于自然数1，2，3，……的基本理论。当哥德尔开始思考希尔伯特纲领时，希尔伯特的学生阿克曼和冯·诺伊曼似乎正朝着用有限性方法证明PA的一致性的方向大步迈进。他们二人都已经为PA的一个受限的子系统找到了这样的证明。且被认为，只要克服一些技术上的困难，成功将指日可待。哥德尔本人很可能也持这种看法。无论如何，他决定去证明那些较之PA更强的系统的一致性。由于已经有了一些重要的关于相对一致性的证明，所以这是一个很自然的想法。哥德尔曾经希望用有限性方法把这种可以包含实数算术的更强的系统的一致性还原为PA的一致性。这在很大程度上是在遵循希尔伯特的路线：希尔伯特曾经把欧几里得几何的一致性还原为实数算术的一致性，现在哥德尔希望把这种还原继续进行下去。如果哥德尔成功了，那么希尔伯特的追随者们对PA的一致性的证明就将自动为实数算术的一致性提供证明，这样就满足了希尔伯特在1900年提出的第二个问题的要求。然而这是做不到的。哥德尔不仅在这项努力中失败了，他还证明了他是不可能成功的。最终，他非但没能像他曾经希望地那样去抵挡布劳威尔—外尔的批评来帮助保卫数学，反而有力地葬送了希尔伯特纲领。

当哥德尔开始思考这些问题时，他重新思考了从外部而不是从内部考察一个形式逻辑系统的意思。罗素和怀特海已经相当令人信服地表明，所有的普通数学都可以在这样一个系统内部发展出来。希尔伯特在他的元数学中试图运用严格受限的数学方法去从外部研究这样的系统。那么，为什么元数学自身不能在一个形式逻辑系统内部发展出来呢？从外部看，这些系统包含着符号串之间的关系。从内部看，这些系统能够表达关于不同数学对象（包括自然数）的命题。而且，想到用自然数来为符号串编码的方式并不困难。啊哈！通过使用这样的代码，外部就可以被带到内部了。为了说明这些代码的用处，让我们再来看看“任何一个恋爱中的人都是快乐的”这个前提是如何被符号化的：

$$(\forall x)((\exists y)\, \boldsymbol{L}(x, y) \supset \boldsymbol{H}(x)) \quad (*)$$

我们这里所有的只是10个符号的一个排列或串，这10个符号是：

$$, \boldsymbol{L} \boldsymbol{H} \supset \forall \exists \boldsymbol{x} \boldsymbol{y} (\ )$$

我们可以采用一种简单的编码方案，其中每一个符号都被一个十进制的阿拉伯数字所替代，例如：

,	$\boldsymbol{L}$	$\boldsymbol{H}$	$\supset$	$\forall$	$\exists$	$\boldsymbol{x}$	$\boldsymbol{y}$	(	)
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
0	1	2	3	4	5	6	7	8	9

对于（*），用上面所示的数字代替那些符号，我们便得到了编码数

846988579186079328699.

请注意，不仅从符号串到它的编码数是容易的，反过来也同样容易。当然，如果符号超过了10个，我们就必须使用一种不同的编码方式，但这并不会带来什么困难。例如，倘若给每一个符号配上两个十进位的数字，那么我们就可以为多达100个的符号编码。从本质上讲，同样的方法可以被用于任何形式逻辑系统。因此，这类系统（所有这些系统从外部看都呈现为符号串）的不同表达方式都可以用自然数来编码。[9]

如何在那些极其相似的系统中用编码去发展出形式逻辑系统的元数学，这对哥德尔来说是不成问题的。但在实际过程中，根据维也纳学派内部广为传播的那些律令，他发现自己的想法是被严格禁止的。哥德尔发现存在着这样的命题，它们从系统外部看是真命题，但却无法在系统内部获得证明。在维也纳学派的许多支持者看来，除了可证明性，数学真理性的任何其他观念都是无意义的，都只是唯心论形而上学的怪胎。哥德尔没有受这些信念影响，他得到了一个完全相反的非凡结论：一种有意义的数学真理的观念不仅是存在的，而且其范围还超出了任何给定

的形式系统的证明能力。这个结论适用于相当广泛的形式逻辑系统，从像PA那样的相对较弱的系统到怀特海和罗素的《数学原理》（PM）那样的囊括了全部古典数学的系统，对于这些系统中的任何一个，都存在着该系统中可表达而不可证的真命题。在哥德尔1931年发表的那篇卓越的论文——“论《数学原理》及有关系统的形式不可判定命题”中，他选择对形式系统PM给出了他的结果，从而说明即使很强逻辑的系统也不可能把全部数学真理包含在内。[10]

在哥德尔的证明中，关键的一步在于他证明了：一个自然数作为在PM中可证命题的代码数的性质本身可以在PM中表示出来。用此事实，哥德尔可以在PM中构造出一些命题，在那些知道如何使用特定编码的人看来，这些命题可以被看作表达了这样一个断言，即某些命题在PM中是不可证的。也就是说，他能够构造出这样一个命题A，该命题经编码后，可以解释为断言某一命题B在PM中是不可证的。现在，在那些没有获知密码的人看来，命题A只不过是一串符号而已，它表示某个关于自然数的复杂而神秘的命题。但是通过代码，神秘性便消失了：A表示这样一个命题，即某个符号串B表示一个在PM中不可证的命题。A和B通常是不同的命题。哥德尔问，它们是否有可能是相同的呢？事实上，它们可以是相同的。哥德尔能够用他从格奥尔格·康托尔那里学到的一种数学技巧——对角线方法来证明这个结论。运用这种技巧，我们可以使被断言为不可证的命题和那个做出这一断言的命题是同一个命题。换句话说，哥德尔发现了如何获得这样一个非凡的命题，我们将称之为U，它有如下性质：

- U说某个特殊的命题在PM中不可证。
- 那个特殊的命题就是U本身。
- 因此，U说：“U在PM中不可证。”

在维也纳学派内部，人们普遍认为，对于一个类似于PM的系统中所表达的命题来说，命题为“真”的意义就是它根据该系统的规则是可证的。命题U的上述性质使得这一信念不再能够成立了。如果我们承认PM不会撒谎，也就是说在PM中证明的任何命题都是真的，[11] 那么我们就发现U是真的，但它在PM中不可证。如下所示：

1.U是真的。假定它是假的，那么它所表述的内容就是假的，因此它就不可能是不可证的，从而一定是可证的，因此是真的。这就与开始的假定即U是假的相矛盾了。所以它一定是真的。

2.U在PM中是不可证的。既然它是真的，那么它所表述的内容必定为真，所以它在PM中不可证。

3.U的否定（记作¬ U）在PM中不可证。因为U是真的，¬ U必定为假，从而在PM中也不可证。

U具有这样的性质：它和它的否定在PM中都不可证。为了强调这种性质，我们把U称为不可判定命题。但这也不能强调得太过分，因为这种不可判定性只与系统内部的可证明性相关。从我们外部的观点来看，U显然是真的。[12]

现在这里出现了一个难题：我们知道U是真的，尽管它在PM中不可证。既然PM包含了所有的普通数学，那么U为真为什么在PM内部不可证？哥德尔认识到这差一点就是可能的了，但这里存在着一个意想不到的障碍。在PM内部，我们可以证明：

假如**PM**是一致的，那么**U**。

因此，正是PM是一致的这一额外假定，才使U在PM内部得不到证明。既然我们知道U在PM内部是不可证的，我们就必须得出结论说，PM的一致性在PM中不可证。然而，希尔伯特纲领的主旨就在于：用被认为仅仅构成PM的一个非常有限的子集的有限性方法来证明像PM这样的系统的一致性。然而哥德尔却证明了，即使就PM的全部能力而言，它也不足以证明其自身的一致性。因此，至少如最初所设想的那样，希尔伯特纲领走到了尽头！[13]

库尔特·哥德尔，计算机程序设计师

今天，我们知道有这样一种实际的物理装置，它的功能相当于一个通用的信息处理可编程计算机。但在1930年，实现这一切还要再等几十年。不过，熟悉现代程序设计语言的人在看哥德尔那年写的关于不可判定性的论文时，会看到45个编了号的公式序列，它们看起来非常像一个

计算机程序。这种相似决非偶然。在证明作为PM中的一个证明的代码数这种性质可以在PM内部表示出来时，哥德尔不得不解决许多问题，这些问题与那些在设计程序语言和用这些语言编写程序时所面临的问题是相同的。在最基础的层次上，当代的计算机仅能通过由0和1所组成的一些短的字符串来执行简单的基本运算。那些所谓高级程序语言的设计师们面临着这样的任务，他们要向程序员们提供可以包含非常复杂的运算的语句，而且还得使这些程序员们喜欢用它。要想让用这些语句所写出的程序被一台计算机执行，它们就必须被翻译成机器语言——翻译成执行它们所需的详细的基本运算清单。此项工作由特殊的翻译程序来完成，它被称为解释程序或者汇编程序。[27]

在哥德尔对不可判定命题的存在性的证明中，其重点在于这一事实，即在PM中的可证明性可以在PM内部表示出来。哥德尔非常清楚自己要把这一革命性的结果提交给一群极富怀疑精神的读者，他希望能够消除任何疑虑。于是他就面临着这样一个问题，即把与PM的公理和推理规则相对应的字符串的代码数的复杂操作分解开来，并把它们转化成用PM的符号语言所写成的表达式。为了解决这个问题，哥德尔实际上创造了一种特殊的语言，通过这种语言，所需的操作就可以被一步步地展开。[14] 每一步都是由一种对数的运算的定义组成的，经过哥德尔所使用的代码的转换，这种运算与PM表达式的一种平行的运算相对应。这些定义利用了在前面的步骤中已经定义过的项，由哥德尔的特殊语言表达出来。这样设计这种特殊的语言，就可以保证通过这样一种定义而被引入的运算能够在PM内部被恰当地表示出来。

莱布尼茨曾经建议发明这样一种精确的人工语言，能够把人类的大多数思想还原为演算。弗雷格在其《概念文字》中说明了如何能够把握数学家们通常使用的逻辑推理。怀特海和罗素用一种人工逻辑语言成功地发展出了实际的数学。希尔伯特已经建议对这些语言进行元数学的研究。但在哥德尔以前，没有人能够说明如何才能把这些元数学的概念植入这些语言本身。[15]

除了构造出一个不可判定命题U以外，哥德尔还希望证明该命题的陈述并不需要异乎寻常的数学概念。为此，哥德尔使用了初等数论中的一个定理——中国剩余定理。通过该定理，哥德尔说明了所有能够用他的特殊语言表示出来的运算如何能够用自然数算术的基本语言表示出来。[16] 由此可知，不可判定命题U本身也可以用这种基本语言表示出

来。这就意味着，U可以用这样一个字母表写出来，组成这个字母表的变量只能取值任何自然数、算术运算+和×、“=”以及弗雷格逻辑中的基本运算，今天写成¬、∧、∨、∃和∀。这里引人注目的结论是，即使是局限于这个如此贫乏的字母表，在PM内部不可判定的命题也可以被构造出来。

柯尼斯堡会议

1930年8月26日，在维也纳的议会咖啡馆，年仅24岁的库尔特·哥德尔正在与鲁道夫·卡尔纳普谈论着10天后将要在柯尼斯堡召开的“精密科学的认识论会议”。卡尔纳普当时近40岁，他是维也纳学派的领军人物。按照安排，他将发表一个关于数学基础的逻辑主义纲领的演讲，这个纲领在怀特海和罗素的《数学原理》中得到了最充分的体现。卡尔纳普的笔记表明，哥德尔已经把自己的重大发现告诉了他，即《数学原理》中存在着关于自然数的不可判定命题。这两个逻辑学家（以及其他与会者）一起去了柯尼斯堡。会议的第一天共有三个关于数学基础的讲演，每一个都持续1小时。卡尔纳普以逻辑主义为主题开始了他的演讲，异乎寻常的是，他在演讲中只字未提哥德尔的新结果。随后是布劳威尔的一个学生海丁的报告，他讲的是布劳威尔的直觉主义。当天的最后一个讲演是冯·诺依曼做的，他所讲的主题是希尔伯特纲领。[17]

第二天又有三个1小时的演讲，除此之外还有三个20分钟的发言，其中的一个发言是哥德尔做的，主题是他的博士论文，即弗雷格的规则的完备性。会议的第三天有一个关于数学基础的圆桌会议。在会上，哥德尔向众人宣布了他那个惊人的发现。他用了相当长的时间试探性地讨论了从类似PM这样一个系统的一致性证明中可以得到什么。他断言，即使已经知道这样一个系统是一致的，我们也完全有可能在该系统内部证明一个关于自然数的命题，该命题从系统的外部看是假的。因此，一个形式系统的一致性不足以保证在该系统内被证明的命题是正确的。冯·诺伊曼对此表示了赞许，这显然激励哥德尔继续走了下去。哥德尔进而宣称，如果假定像PM这样的系统的一致性，“我们甚至能够给出一些有着简单算术形式的命题的例子”，它们是真的，但在这样一个系统中却是不可证的。“因此”，他继续说道，“假如把这样一个命题的否定式加入”PM，那么我们就可以得到一个一致的系统，但在该系统中有一个假命题是可证的。[18]

冯·诺依曼似乎立刻就理解了哥德尔的工作的意义，他在会议结束时找哥德尔进行了一次讨论。没有任何迹象表明还有其他人认识到发生了什么。冯·诺依曼继续思考这一问题，他确信（原因已在上文解释过了）从哥德尔的结果可以得出一致性本身是不可证的，并进一步断定这就宣告了希尔伯特纲领的失败。当冯·诺依曼写信告诉哥德尔这个信息时，哥德尔已将自己含有同一结论的论文拿去发表了。哥德尔立即做出回复，还寄上了他的文章的预印本。冯·诺依曼回信感谢了哥德尔，并且不无沮丧地说："既然您所建立的一致性的不可证性是您早期的结果的自然延续和深化，那么我当然不会就该主题发表什么了。"[19] 逻辑和数学基础曾经是冯·诺依曼的主要兴趣之一。他成了哥德尔的一个好朋友，曾就哥德尔的工作做了多次演讲，并且称赞哥德尔是自亚里士多德之后最伟大的逻辑学家。[20] 但他自己却终止了逻辑方面的工作。当他10多年后重新对逻辑有了兴趣时，他对逻辑的兴趣体现在了硬件上：通用数字计算机。

冯·诺依曼后来与计算机打交道时的一个合作者曾经讲过这样一个有趣的故事，冯·诺依曼常常谈起他证明算术一致性的努力：

> 一天的工作结束后，[冯·诺依曼]会去上床睡觉，在夜里他经常因为有了新的洞见而醒来……这时他会努力为[算术一致性]给出一个证明，但却未能成功！有一天夜里，他梦见了如何去克服他的困难，并且把他的证明继续推进……第二天一大早，他就着手攻克难关，这一次他又没有成功，又一次在夜里身心疲惫地上床睡了觉，然后开始做梦。这回他又梦见了克服困难的方法，但是当他醒来之后……他看到那里仍然有一道无法跨越的鸿沟。

冯·诺依曼打趣地说，如果他能在第三个夜里梦见它，事情就会是另一种样子了！[21]

哥德尔公布他的惊人发现的那次会议并不是那一周在柯尼斯堡所举行的主要活动，主要活动是德国科学家和医生协会在柯尼斯堡举行的大会。在圆桌会议的第二天，大卫·希尔伯特就发表了开幕演说。正是在这里，希尔伯特明确地喊出了他的口号："wir müssen wissen；wir werden wissen"（我们必须知道，我们将会知道），他相信所有数学问题都必须得到回答，而且将会得到回答。这一口号后来刻在了他的墓碑上。哥德尔的不完备性定理表明，如果数学被局限在像PM那样的特殊

的形式系统中，那么希尔伯特的信念就是徒劳的。对于任何一个特定的形式系统，都会有数学问题超越于它。另一方面，从原则上讲，每一个这样的问题都会导向一个更强的系统，在该系统中能够得到这个问题的解答。我们可以想象出更强系统的分层体系，每一个较强的系统都可以解决较弱的系统所遗留的问题。虽然所有这些作为理论问题是无可争议的，但是它在何种程度上会成为数学实践，这个问题却是模糊不清的。哥德尔为数学家们留下了这样一份遗产，即学会用这些更强的系统来解决棘手的问题。虽然一些勇敢的研究者仍然沿着这些思路进行着研究，但大多数数学家仍然不知道这些问题，有些专家也带着极端的怀疑欢迎这项工作。[22]

爱与恨

奥尔伽·陶斯基—托德是哥德尔在维也纳时的同学，她后来成了一位卓越的数论学家。她曾经说，在同学中哥德尔的能力是公认的，每当有同学遇到困难时，哥德尔总是愿意提供帮助。她讲了下面一则有趣的逸事：

> 毫无疑问，哥德尔喜欢班上的异性同学，他自己也从不隐瞒这个事实……有一次，我正在图书馆外面的数学研讨班的小房间里工作。这时门开了，一位个头很小的、非常年轻的姑娘走了进来。她长得很好看……穿着一件美丽的、样式很独特的夏装。没过多久，哥德尔走了进来，于是她站起来，两个人一同离开了。哥德尔似乎很显然是在炫耀。[23]

当哥德尔还是学生的时候，他就遇到了阿黛勒，一个将会成为其终身伴侣的女人。他们交往了十几年才结了婚。当时她是个有夫之妇，在一家夜总会做舞女。[28] 哥德尔的父母不大可能对他的选择感到高兴——这不仅是因为她比哥德尔大6岁，而且还因为她是一个罗马天主教徒。而且，维也纳的舞女或多或少都有着不好的名声，人们认为她们会为不多的一笔钱而提供性服务。[24] 也许正是由于这些因素，哥德尔在处理自己和阿黛勒的关系上相当慎重，他们两人在婚前似乎保持着亲密的私人关系。当他们最终结婚时，阿黛勒的存在令哥德尔的同事们感到相当惊讶。[25] 鲁道夫（他一直单身）在他的弟弟死后不久写道："对于我弟弟的婚姻，我将不会擅自做出评价。"[26] 婚姻的幸福永远是一个巨

大的秘密，老人和所谓的智者对此所做出的预言往往会出错。哥德尔的婚姻就是这样，事实证明，他和阿黛勒的婚姻生活是持久而快乐的。

当哥德尔试图在奥地利开始他的职业生涯时，这个国家正陷于政治、社会和经济的灾难性的动荡之中。当第一次世界大战即将结束之时，这个在奥匈帝国的废墟上建立起来的操德语的国家被协约国禁止去做大多数奥地利人想做的事情：与德国统一成一个国家。无论如何，独立民主的奥地利并没有存在多久。1927年，在法西斯的武装党卫队和社会民主党的防御同盟之间所展开的一场强度不大的内战达到了白热化：当一个老人和一个小孩被反动分子所杀害，而一个陪审员拒绝宣判杀人犯们有罪时，社会民主党人组织了一场规模浩大的示威游行。在这场游行中，司法部大楼被焚烧，大约有一百多人丧生。1929年底，共和国的总统已经通过紧急事态法令而掌握了统治权。同时，席卷世界的严重的经济危机（在美国称为“大萧条”）使得保持克制再也不可能了。1932年被选举上台的多尔福斯政权成了独裁的政府，它终止了议会的任何有意义的职能。原本就已经很坏的事态，现在变得更糟了。到了1934年的早些时候，希特勒已经在德国掌权，除多尔福斯的祖国阵线以外，所有其他的政党都被废除了。几个月后，多尔福斯在奥地利纳粹的一次未遂的政变中被谋杀。他的继任者舒施尼希在墨索里尼的帮助下，没有让希特勒染指奥地利。但好景不长，到了1938年3月，奥地利终于被纳粹德国所吞并。

1933年2月，哥德尔获得官方任命当上了讲师，从此开始了他那漫长的学术升迁过程。与此同时，他积极参加了数学家卡尔·门格尔（他也是维也纳学派中的一个积极分子）所主持的一个讨论会，以及他的论文指导老师哈恩所主持的一个逻辑研讨班。哥德尔的一大批有趣的成果都是出自这一时期，其中一些重要成果是在门格尔的讨论会文集中作为小文章发表的。[27] 1933年夏，哥德尔在困难的条件下开设了他任讲师之后的第一门课。曾经有一周，维也纳有好几个地方都有纳粹恐怖分子安放的炸弹爆炸。由于纳粹的活动，大学不得不在某一天关闭了。

在这种情况下，当新成立的普林斯顿高等研究院邀请哥德尔去做1933—1934学年的访问研究时，哥德尔是很难拒绝的。他不仅可以躲过国内政治的疯狂，而且他也很期望能与阿尔伯特·爱因斯坦和冯·诺依曼这样的科学巨星共事。然而，离开家人和朋友（也许主要是阿黛勒）几乎一年的时间，肯定会给这个害羞的患有疑病症的年轻人带来了不少忧

虑。事实上，当哥德尔去乘坐那条将带他穿越大西洋的渡轮时，他确定自己得了感冒，又打道回府了。只是在家人的劝说下，他才乘了另一艘船做了此次远航。

哥德尔在普林斯顿这一年的生活几乎不为人所知。他于当年12月在马萨诸塞的剑桥所做的一个演讲的演讲稿被保存下来了。此外，第2年春天他在普林斯顿所做的一系列演讲的稿子也肯定被保留下来了，但关于他个人生活的信息我们已经无从获得。我们所知道的只是当年6月他返回了维也纳，此后没过几个月，他就精神崩溃了。他在普克斯多夫疗养院住了一段时间，“该疗养院是专为富人们开设的，那里有温泉疗养区，设有门诊部，还有病人休养所”。在那里，给他治疗的是诺贝尔奖获得者——精神病医生尤利乌斯·瓦格纳—尧雷格。[28] 当哥德尔返回奥地利时，奥地利遭遇了许多不幸的事件。7月末，多尔福斯在纳粹的一次未遂政变中被暗杀；前一天，哥德尔的博士论文导师汉斯·哈恩死于癌症手术的并发症；大学里的情况日益恶化，行政负责人被要求加入法西斯的祖国阵线；许多被认为是左派的教授，甚至那些从不关心政治的犹太学者们，都被解除了职务。不过，如今已经不可能知道这些事件对哥德尔的精神崩溃起过什么样的作用。

回想起来，法西斯主义的稳步发展中所包含的危险是很容易看出来的。但对那些有能力预见未来而选择逃走的人来说，事情并不那么简单，他们认为事情总会得到解决。哥德尔的哥哥写道，他们家里没有一个人“对政治感兴趣”，因此，他们没能理解1933年希特勒在德国掌权意味着什么。然而，他继续写道：

> 有两个事件迅速地擦亮了我们的眼睛：总理多尔福斯被暗杀，以及哲学家石里克被暗杀。石里克教授是被一个国家社会主义的学生暗害的，我的哥哥曾经参加过他的学派。[29]

哥德尔一方面保持着与普林斯顿高等研究院的联系，从而为将来准备更多的可能性；另一方面，他继续在维也纳发展自己的学术事业。1935年5月，他在大学里开设了他的第二门课程。同年9月，他又一次到普林斯顿高等研究院做访问学者，这一次他没有在美国待多久。由于精神极度沮丧，他辞去了自己的职位，12月初就回到了家。哥德尔后来曾说，石里克被刺杀的1936年是他一生中最糟糕的年头。他的精神状况越来越差，大多数时间都在疗养院度过。但1937年是一个巨大的转机。这

年6月，当哥德尔在大学里教授一门集合论课程时，他关于康托尔的连续统假设的工作取得了一个重大突破。连续统假设在希尔伯特1900年著名的问题列表中被列为第一个问题。（后面我们会具体讲到。）

1938年3月，希特勒入侵奥地利，并将其并入德国。同年10月，哥德尔第三次去美国访问。此时距他和阿黛勒结婚不过2周，他把阿黛勒留在了维也纳。[29] 他在美国度过的这一年可谓成果颇丰。他在普林斯顿度过了秋季学期，并在那里做了关于康托尔的连续统假设的发现的演讲，其后的春季学期他去了圣母大学做访问研究。他的老同事卡尔·门格尔在逃出维也纳之后，就在该大学任教。但是当整个学年结束后，1939年6月末，哥德尔又返回维也纳与阿黛勒团聚。2个月后，德国就入侵了波兰，从而挑起了第二次世界大战。

此时，哥德尔所处的维也纳已经是纳粹德国的一部分，它正在作为希特勒“新秩序”的一部分而被系统地改造。在大学里，讲师的职位已经被废除，取而代之的则是一种被称为“新秩序讲师”（*Dozent neuer Ordnung* ）的新的职位。这个新职位会提供少量薪水，但它需要重新申请，而且候选人必须经过政治观点和种族纯洁性的考察。当年9月，也就是第二次世界大战爆发后不久，哥德尔就提出了申请。令他惊讶和气愤的是，他的申请竟然没有被批准。负责讲师申请的官吏在呈交给院长的报告书中指出，哥德尔曾在“犹太教授哈恩”的指导下工作过，而且他还出入于“犹太人—的自由主义的”圈子。另一方面，也从未听说他有过任何“反对国家社会主义”的言论。在这种情况下，无论是批准还是拒绝申请都是不可能的。一个不可判定命题！

耽搁了几个月之后，哥德尔又遭受了另一个严重打击。他被要求参加体格检查，以确定他是否适合服兵役。令他惊讶的事件又一次发生了，他被宣布为符合驻防的条件。差不多就在这个时候，哥德尔和阿黛勒于11月从他们在郊外租的房子里搬出，迁入了他们新近在城里买的一套公寓。[30] 哥德尔明显地忘记了发生在他周围的一切，这只能被解释为一种病态的拒绝。犹太人古斯塔夫·伯格曼是维也纳学派的一个成员，1938年10月，他随着一股犹太难民潮到了美国，他曾经讲过一个故事，很好地说明了这一点。当伯格曼在美国安定下来后不久，哥德尔邀他共进午餐（然后在普林斯顿访问）。哥德尔问：“是什么让你决定到美国来的，伯格曼先生？”这使他大吃一惊。[31] 使他最终认识到哥德尔的危险境况的事情似乎是，他搬家之后不久，他在街上遭到一伙流氓殴

打，眼镜也被打掉了。[32]

在德国闪电般地占领波兰之后，1939—1940年的冬天是有名的“假战”（Phony war）时期。此时距德军攻入西欧，并且导致法国的溃败还有几个月。直到1941年的6月，德国才开始进攻苏联。事实上，德国和苏联曾经签订有互不侵犯条约。此前苏联一直都在向德国提供军用物资。1939年12月，哥德尔终于决定尽一切可能离开欧洲。为此，他和阿黛勒需要获得德国政府的出境许可证以及美国的签证。这两件事没有一件是容易办到的。在这方面，普林斯顿高等研究院的新任院长弗兰克·艾德洛特功不可没。在与美国国务院交涉时，他有意把真实情况做了一点更改。虽然他很清楚哥德尔还不是教授，但在他的信件中，哥德尔成了“哥德尔教授”。在回答哥德尔在普林斯顿高等研究院的教学职责是什么的时候，他很冷静地撒谎说：“哥德尔教授的职责将包括教学”，不过是在一个高级的因而是非正式的地位上。此外，艾德洛特还在给华盛顿的德国大使馆写的信中强调，哥德尔是一个“雅利安人”，而且是世界上最伟大的数学家之一。这个计谋奏效了，所有必需的文件均已办妥，哥德尔现在可以离开欧洲了。然而，人们认为穿越大西洋实在是太危险了，于是他们绕了一大圈穿过西伯利亚到达日本，然后横渡太平洋，最后乘火车抵达普林斯顿。当他们到达目的地时，已经是3月中旬了。[33]

奥斯卡·摩根斯滕是最先去迎接哥德尔的人之一，他后来成了哥德尔的一个好朋友。摩根斯滕是一个经济学家，他在维也纳学派偶然认识了哥德尔。当他从领导职位上被解聘之后，他接受了普林斯顿大学的一个教授职位。当他迫切地向哥德尔询问维也纳当前的形势时，哥德尔说：“咖啡很糟糕。”[34] 这个回答令他倒吸一口冷气。

希尔伯特的宣言

在希尔伯特1900年的问题列表中，排在首位的就是康托尔的连续统假设。它断言由实数组成的无穷集合只有两种大小。小的无穷集合是那些与自然数集同样大的集合，也就是说这些集合能够与集合｛1，2，3，……｝一一对应起来。大的无穷集合是那些能够与实数集一一对应起来的集合。连续统假设说，任何一个由实数组成的无穷集合必定是这两种类型当中的一种，因此决不存在一个实数无穷集，它的大小介于这两者之间。（用康托尔无穷基数的语言来表达，就是任何一个由实数组

成的无穷集的基数要么是$\aleph_0$，要么是C。）希尔伯特在他的讲演中称，连续统假设“似乎非常真实”，但“尽管付出了最为艰辛的努力，还没有人能够成功地证明”它。[35] 四分之一个世纪以后，希尔伯特又回到了这个问题，他宣称自己能够用他的元数学方法来证明连续统假设。然而，事实证明这不过是一个幻觉。1934年，波兰数学家瓦茨拉夫·谢尔品斯基发表了一篇论文，这篇文章专门讨论了业已发现的与连续统假设相等同或相关的命题。然而，尽管有所有这些持续不断的“艰辛的努力”，连续统假设是否为真仍然不能判定。

哥德尔开始相信，对于能够充当数学基础的现有形式系统而言，连续统假设是不可判定的。这样的系统不仅包括罗素和怀特海的PM，而且也包括建立在集合论公理的基础之上的系统。但他只能部分地证明他的信念是合理的：1937年，他看出了如何证明连续统假设在这些系统中是不可能被否证的。[36] 虽然他确信连续统假设同样也不可能在这些系统中得到证明，但他从未能够证明这一点。（四分之一个世纪以后，哥德尔被证明是正确的，保罗·科恩发展出了强有力的新方法，并借此证明了连续统假设在所讨论的系统中确实是不可判定的。）

希尔伯特在1900年的巴黎演讲中，以及1930年的柯尼斯堡退休演讲中，都宣布了他对于任何一个数学问题的可解性的信念。但数学家们一直无法解决康托尔的连续统问题，这是否意味着希尔伯特错了？哥德尔已发现涉及自然数的不可判定命题在所讨论的形式系统内部是不可判定的，但正如我们已经看到的那样，从外部来看，它们显然是真的。不过连续统假设就不同了：哥德尔的工作没有暗示它到底是真是假。直到这时，哥德尔都未受狭隘的基础观点的阻碍，他能够运用任何所需要的数学方法开路前行。但是现在，他的结果强迫他停下来去思考他的工作的哲学含义。

数学家们通常处理的单个的实数，比如π和$\sqrt{2}$，可以在像PM这样的形式系统中被定义。但人们在康托尔的时代就已经很清楚，在这样的系统中，由所有能被定义的数所组成的集合的基数是$\aleph_0$，而由全部实数所组成的集合的基数是。我们知道后者更大。因此，大多数实数没有定义：它们是不可判定的。这是很奇怪的，你如何能够数出那些你不能定义的东西？谈论一个其中有许多数字不可定义的实数集，这有意义吗？也许连续统假设的不可判定性（哥德尔做此猜想，后来保罗·科恩给出了证明）告诉我们，它并没有一个清楚的意义，它本来就是含糊不清

的。解决这个问题，就是顽强地面对实无限在数学中扮演着什么角色的问题。弗雷格曾经断言这个问题将导致一场“关系重大的、决定性的斗争”。[37]

当哥德尔得到了关于连续统假设的研究结果后不久，他就此做了一些演讲。从演讲的手稿中可以看出，他的态度是模棱两可的。他暗示连续统假设很可能是“绝对不可判定的”，这说明希尔伯特相信任何一个数学问题都可以被解决是错误的。到了20世纪40年代的早期，哥德尔转向了哲学研究，毫无疑问，这在一定程度上有助于他思考自己对无穷集合的看法。他尤其对莱布尼茨情有独钟，这位古典哲学家对他最有亲和力。

高等研究院的成员没有做演讲、指导学生和发表文章的义务，在这种宽松的环境下，哥德尔只有在收到非常特殊的邀请时才会做演讲或是发表文章。发出这样的邀请的主要是《在世哲学家文库》（*Library of Living Philosophers*）。它是一个丛书系列，每一本只写一位在世的哲学家。每一卷都是由对该哲学家思想的评论文章所组成，而哲学家本人的回答也附在后面。哥德尔应邀评论了伯特兰·罗素、阿尔伯特·爱因斯坦和鲁道夫·卡尔纳普。罗素卷在1944年出版，其中有一篇哥德尔写的相当惊人的文章。在对罗素的数理逻辑做了深刻的讨论之后，哥德尔宣称：集合与概念可以被“认为是真实的对象……它们独立于我们的定义和构造而存在……假定这样的对象与假定物理对象具有同样的合法性，我们有相当多的理由相信它们存在”。这说得多含糊呀！3年后，哥德尔应邀写了一篇解释连续统假设的文章。他在文中重申了他对集合真实存在的信念，强调了现存的形式系统必定是不完备的和能够扩展的，并且预测人们将会找到新的公理，从而通过证明连续统假设是错的来最终解决连续统假设的问题。[38]

在哥德尔研究连续统假设之前，他在与哲学问题打交道时，往往忽略了那些会妨碍他看到对他来说是一清二楚的东西的其他哲学问题。但是现在，他已经深深地沉浸在了哲学之中。数究竟是什么？它们仅仅是人类的构造，还是有着某种客观的存在性？在地球上的人们断定2+2=4之前，它是真的吗？这些问题已经被人们争论了许多个世纪。有一种学说认为，抽象的对象（比如数和数的集合）是客观存在的，人们只能发现而不是发明它们的属性。这种学说通常被归之于柏拉图，因此也被称为柏拉图主义。哥德尔坚持这种学说，标志着他的观点发生了明确转

向。1933年，他在马萨诸塞州的剑桥做了一次演讲，那时他还宣称柏拉图主义不可能“让任何批判性的心灵感到满意”。[39] 在20世纪的最后几十年里，集合论的研究者们一直在按照哥德尔的指令来寻求新的公理，但是尽管有许多有趣的工作，连续统假设仍然没能得到解决。

在哥德尔为罗素卷所写的文章中，最令人惊讶的一段话涉及莱布尼茨关于一种普遍文字的心爱的计划。在莱布尼茨去世2个世纪之后，哥德尔也希望这样一种语言能够被发展出来，它将使数学实践发生革命：

> 没有必要放弃希望。如果我们相信莱布尼茨所说的，他已经在相当程度上发展了这种推理演算，但要等到种子落在肥沃的土地上时才会公布它，那么他关于普遍文字的文章就并非是在讨论一个乌托邦式的计划。他甚至还估计了几位优秀的科学家发展出他的这种演算所需要的时间，那时“人类将会拥有一种新的工具，它对理性力量的增强将远远超过任何光学工具对视觉力量的增强。”他认为这段时间是5年，并且宣称学习他的方法并不比学习他那个时代的数学或哲学更为困难。[40]

我们已经看到，与布尔和弗雷格后来的成果相比，莱布尼茨通过一种推理演算所产生的东西是微不足道的，尽管这在他那个时代是令人惊异的。那么，哥德尔在想些什么呢？唉，他似乎相信有一个阴谋抑制住了莱布尼茨的思想。哥德尔在许多方面都有着非常古怪的信念，事实上，他至少是有一点妄想狂的症状。然而，他在逻辑学家中的威望是如此之高，以至于逻辑学家很难轻易摒弃他的任何想法。（我们在后面会更多地谈到哥德尔的精神问题。）

当哥德尔应《在世哲学家文库》之邀写作关于爱因斯坦的文章时，他选择爱因斯坦的相对论与康德哲学之间的关系作为文章的主题。他发现广义相对论（爱因斯坦的引力理论）的方程有一个解与物理学家们曾经设想过的非常不同。引人注目的是，哥德尔所找到的方程的解表示这样一个宇宙，只要时间足够长，速度足够快，在该宇宙中的旅行就可以回到过去。很自然地，这样一个世界容易受到时间旅行悖论的攻击，科幻小说的读者对此是很熟悉的：例如，一个人能够回到过去杀死他童年时的祖父母吗？针对这个难题，哥德尔提出了一种令人惊讶的非哲学的解答。他指出，由于做此种旅行所需要的燃料数量巨大，所以这样一种旅行是不切实际的。

在发表文章以前，哥德尔通常要一遍又一遍地细心审订，直至令自己完全满意为止。即使在发表之后，他也会利用重印的机会进行进一步的修订。眼瞅着截稿日期一拖再拖，他的编辑们常常因此而感到灰心丧气。哥德尔曾经许诺为《在世哲学家文库》写一篇关于鲁道夫·卡尔纳普的文章，最后出版的卡尔纳普卷并没有收入他的这篇文章。然而，人们在哥德尔的论文中发现了六篇他特意为批判卡尔纳普关于逻辑和数学的观点而写的文章。后来，他的《文集》的编辑决定发表其中的两篇。在他的论文中，人们还发现了他在1951年圣诞节期间在罗得岛的普罗维登斯所做的一个讲演的草稿（里面有各种插入语、删掉的部分和脚注）。[30] 讲演的题目是“关于数学基础的一些基本定理及其蕴含”。在这个讲演中，哥德尔实际上是把希尔伯特关于任何一个数学问题的可解性的宣言放在了人的心灵本性的背景之下。哥德尔提出了这样一个问题，即人的心灵从本质上讲是否等同于一台计算机。直到现在，当人们谈到人工智能的前景时，仍要对这个问题进行激烈的争论。哥德尔本人并没有给出这个问题的答案（尽管人们后来知道，哥德尔相信正确的答案是否定的），他只是说任何一个答案都是“与唯物论哲学截然对立的”。假如人的心灵的全部能力都能被一台有限的机械装置模仿出来，那么哥德尔本人的不完备性定理就可以被用来说明，某个关于自然数的命题虽然是真的，却不能被人类证明，它是一个绝对不可判定的命题。这显然与希尔伯特的宣言相矛盾。但是按照哥德尔的说法，它也需要一些唯心论哲学的手段，为的是让一个假定自然数（其属性超出了人类心灵所能探知的程度）客观存在的陈述有意义。而另一方面，哥德尔推理说，如果人的心灵不能被还原为机械装置，而物理的大脑却可以被如此还原（他认为这是显然的），那么就说明心灵超越了物理实在，这同样与唯物论不相容。与其说这则论证完全令人信服，不如说它是把理论逻辑、人类生理学、计算机的最终潜力以及基础哲学放在一起加以考虑。哥德尔再次展示了他那令人惊叹的独辟蹊径思考问题的能力。[41]

一个奇特的人和一个悲哀的结局

当库尔特·哥德尔接近退休年龄时，他希望当时正在耶鲁大学执教的逻辑学家亚伯拉罕·鲁宾逊能够接任他在普林斯顿高等研究院的职位。但这还没来得及成为现实，鲁宾逊经诊断患上了不宜动手术的胰腺癌，没过多久就去世了。在他去世前的最后几个月里，鲁宾逊收到了哥德尔这样的一封信：

> 考虑到去年（此时鲁宾逊在普林斯顿高等研究院待了一段时间）在我们的讨论中我所说的话，你能想象我对于你的病感到多么遗憾，这不仅是我个人的看法，而且也是逻辑学和高等研究院的遗憾。
>
> 如你所知，我对于许多事情都有着非正统的看法。其中有两点在这里适用：
>
> 1. 我不相信任何医学诊断是百分之百确定的。
>
> 2. 在我看来，说我们的自我是由蛋白质分子构成的，这是人们所做过的最可笑的论断之一。
>
> 我希望你至少同意第二个观点。我很高兴地听说，尽管你病了，但你还是能够在数学系里度过一段时光。我相信这将给你带来一些愉快的消遣。[42]

这封信具有典型的哥德尔风格。他所说的自己不相信医学诊断当然是一种保守的说法。当他因前列腺肥大导致输尿管阻塞而痛苦不堪时，他不仅拒绝接受诊断，而且还坚持说，只要加大他轻泻剂的剂量，他的问题就可解决。那时，他对于轻泻剂已经是相当依赖了。他曾愤怒地拔出插在他体内的导尿管。由于拒绝进行通常能够缓解这种病症的手术，他最终接受了导尿管，并且一直使用到去世。他试图通过间接地谈论自己关于心灵不只是蛋白质分子的信念而去安抚鲁宾逊。显然，这暗示了人是有来生的。这是哥德尔的另一种典型风格。

在哥德尔的非正统观点与彻底的妄想症之间的界限并不总是清楚的。当摩根斯滕发现哥德尔极为严肃地看待幽灵时，他记录了自己当时的惊讶。更为离奇的是，哥德尔确信在他普林斯顿的公寓里，冰箱和电暖炉正在释放有毒气体，为此他和阿黛勒搬了好几次家。最终，他径直搬走了那些讨厌的家用电器，于是他的公寓成了“一个在冬天很不舒服的地方”。

当哥德尔试图成为一个美国公民时，他以一种典型的哥德尔风格为他在法官那儿的面试做准备。通常，法官都是马马虎虎地检查一下当事人关于美国宪法的知识，但哥德尔却对美国宪法做了那种只有他才会做出的细致分析，而且当他得出结论说美国宪法确实是不一致的时候，他

变得异常激动。摩根斯滕和爱因斯坦是哥德尔的入籍证人，当他们三人开车去新泽西州首府特伦顿办理手续时，他们试着不让哥德尔去想他的发现，因为他们害怕哥德尔在面试时谈起它会给入籍带来麻烦。爱因斯坦讲了一个又一个的笑话，但是当法官问哥德尔是否认为美国会出现像德国那样的一个独裁政府时，哥德尔便开始解释他的发现。幸运的是，法官很快就知道了他在与一个什么样的人打交道，所以他打断了哥德尔的讲话。这样，最终还是皆大欢喜了。

任何人听了这些关于哥德尔古怪之处的逸事，都可能忍俊不禁。但并不是所有的逸事都是那么有趣的。哥德尔后来陷入了一种怀疑食物安全性的妄想症之中，他深爱的妻子也因为病重而不能给他太多的帮助，此时他竟绝食而死。就这样，1978年1月14日，20世纪最伟大的思想家之一离我们而去了。

第7章　图灵构想通用计算机

早在1834年，查尔斯·巴贝奇就构想出了一种自动的计算机器，这就是他虽然提出但从未制造出来的分析机，旨在完成各种数值计算。[31] 为了强调他的机器的威力和应用范围，巴贝奇开玩笑说："除了编乡村舞蹈，它可以做任何事情。"[1] 尽管在巴贝奇看来，为了计算的目的而设计出来的机器显然是不可能编舞的，但我们今天却认为实现这些完全没有问题。事实上，今天的计算机完全可以通过程序设计而编出乡村舞蹈（尽管质量可能不是最高的）。今天，如果我们想用一个类似的比喻来强调计算机的威力和应用范围，那么我们就会发现，写出这个句子并不容易。几乎任何可以设想的包含符号、数字或文本的任务都或者已经处于计算机的能力范围之内，或者某位专家预言不久就可以做到。显然，我们关于计算的概念已经发生了天翻地覆的变化。1935年，阿兰·图灵在解决大卫·希尔伯特所提出的一个数理逻辑问题的过程中，提出了基本的概念。

巴贝奇原本打算完全用类似齿轮这样的机械部件制造出他的机器，毫不奇怪，由于这个装置的复杂性，他没有成功。只是到了20世纪30年代，随着使用电子继电器的机电计算器的发展，应用范围达到了巴贝奇所设想的程度的机器才得以制造出来。但是在30至40年代，没有一个参与这项工作的人曾经谈起过能够超越于数学演算的机器。正如我们将会看到的，第一个使巴贝奇的成为现实的人是霍华德·艾肯。他写道：

> 如果，一台为了微分方程数值解而设计的机器与百货商店里的一台用来开账单的机器的基本逻辑是一致的，那么我将把这看成我所遇到过的最令人惊异的一致。[2]

艾肯是在1956年说这段话的，那时计算机通过编程来实现这两项任务在商业上已经是可能的了。如果艾肯认识到阿兰·图灵20年前所发表的论文的重要性，他就不会做出如此可笑的论断了。

帝国的孩子

阿兰·图灵的父亲尤利乌斯·图灵是一个在印度工作的有突出业绩的公务员。1907年春，在工作了10年之后，他准备回到英国去。当驶往故乡的船通过太平洋时，他遇上了阿兰的母亲埃塞尔·萨拉·斯通尼。埃塞尔·萨拉出生于马德拉斯，她在爱尔兰长大，在巴黎也待过一段时间，那时已经回过印度。在归途中，两人堕入了情网，他们最终一起去了美国的黄石公园旅行。经她父亲同意，他们在返回印度的那年秋天在都柏林结婚。阿兰的哥哥约翰出生于1908年9月。尤利乌斯的工作使得他需要经常在印度南部旅行，埃塞尔·萨拉和这个婴儿总是陪伴在他的身旁。1911年秋，埃塞尔·萨拉在旅行途中怀上了阿兰，在尤利乌斯设法又一次被获准离开之后，他们一同乘船到了英国。阿兰·图灵于1912年6月23日出生于伦敦。[3]



（阿兰·图灵；蒙伦敦国立肖像陈列馆惠允）

对于图灵的家庭来说，帝国的无情现实使生活变得十分艰难。父亲在印度工作，在那里热带病特别容易感染年幼的孩子，而且他们也接受不到适当的教育。母亲或者是和丈夫在一起，或者是和孩子们在一起；只有当父亲休假时，她才可以与所有人团聚。当阿兰15个月大的时候，他的母亲把他和4岁大的哥哥留在了英国，让他们住在一个退伍上校家里，自己则回到了印度。图灵太太在1915年与孩子们一起待了几个月，1916年春，父母都回到了家，但这时德国潜艇成了另一个潜藏着的危险，于是当丈夫返回印度时，图灵太太留在了英国。事实上，正是由于严酷的战争，母亲才得以陪伴在阿兰的身旁。他是一个早熟的快乐的孩子，很乐意交朋友，同时也是一个做事笨手笨脚的邋遢孩子。虽然许多6岁的孩子都离开家住在学校里，但阿兰的母亲一直让他与自己待在一起，并把他送到一所当地的日间学校去学习正规教育所不可或缺的拉丁语，发刮嚓声的钢笔，漏水的自来水笔以及阿兰非常潦草的笔迹使得这些课程对于阿兰来说一点都不轻松。

1919年，当母亲再次到印度去的时候，7岁的阿兰又回到了上校的家。时隔不到2年，当图灵太太回来时，她发现她的孩子并没有茁壮成长。她离开时的那个快乐的小男孩消失了踪影，她所看到的是一个不合群的、性格内向的孩子，他的基础教育已经被严重地忽视了。在尽可能地做了努力之后，她把阿兰送到了他哥哥约翰所在的寄宿学校。兄弟俩只在一起待了几个月，约翰就去读一所公立学校了。[32] 于是，当暑假过后父母离开时，阿兰只得独自过寄宿生活。他可怜地追着父母远去的汽车，以此来表达自己对前景的感受。

到了14岁时，阿兰·图灵开始住在舍波恩（Sherborne）公立学校，他对科学和数学的激情已经被点燃。他发现自己所处的环境特别重视竞技体育，而对数学特别的轻视。他的一位老师曾认为科学是“低等的和狡猾的”，并说数学在屋子里散发出一种难闻的味道。[4] 阿兰的数学天赋虽然得到了认可，却并没有受到重视，他的父母被警告说，他有可能只会成为一个科学专家。最重要的是，他的作业污迹斑斑，字体几乎无法辨认。与此同时，阿兰与其他孩子并不怎么交往，他对所上课程毫不关心（但考试却考得很好），自己做着零碎的数学研究，还学习了爱因斯坦的相对论。

当阿兰遇到他的朋友克里斯托弗·毛肯时，他的生活发生了改变。克里斯托弗和阿兰都喜欢科学和数学，但与阿兰不同，克里斯托弗是一个勤奋的学生，他认真听所有课程，作业也写得极为整洁。阿兰对克里斯托弗的羡慕无以复加，并决心向他学习。不知是在什么时候，阿兰·图灵知道了自己的同性恋倾向，但可以很自然地猜测，至少对于阿兰来说，他与毛肯的友谊带有性的色彩。图灵的传记作者把阿兰的情感称之为“初恋”，其强烈程度也确实是够了。如果不是因为悲剧发生，没有人会知道这种关系会如何发展，阿兰的感情会怎样调整。阿兰不知道，他的朋友一直患着结核病。他死于1930年2月，作为一个完美的符号永远铭刻在了图灵的心中。[5]

在舍波恩的最后1年，图灵的学习变得非常优秀，他得到了剑桥大学国王学院的一份奖学金。除了提供食宿，每年他还可获得80英镑的津贴，当时差不多抵得上一个熟练的工人年薪的一半。[6] 虽然数学在舍波恩的名声不佳，但在剑桥大学，图灵感到这里的氛围可以使他的数学天才充分施展出来。G.H.哈代（1877—1947）是剑桥的大数学家，他于1908年出版的《纯粹数学教程》（*Course of Pure Mathematics* ）已经成了一本经典教科书，一代又一代的青年数学家从中掌握了极限过程的基本性质。（现在仍然在销售）有一个关于来自马德拉斯的自学成才的邮政职员拉马努金的公众电视节目（1988年首播）对哈代进行了描绘，节目中说，哈代使拉马努金的数学天才大放异彩。

图灵可以上哈代以及数学物理学家和天文学家阿瑟·爱丁顿爵士的课。1919年，爱丁顿曾带领远征队抵达西非，那里的日全食使得对光线经过太阳附近时的行为进行观测成为可能，于是第一次证实了爱因斯坦在其广义相对论中关于太阳引力使光线偏折的预言。爱丁顿在课上提出了这样一个问题，为什么如此众多的统计观测似乎都沿着著名的钟形正态分布曲线排列。他的课还涵盖了使物理学发生革命的当时非常新的量子理论。但是在这个领域，真正引起图灵注意的是一本新近出版的约翰·冯·诺依曼写的书，这本论述量子力学的数学基础的书是阿兰在舍波恩获的奖。

爱丁顿在课堂上强调的无所不在的钟形正态分布曲线使图灵入了迷，他试图找到其背后的数学解释。他证明了大量统计分布的确在极限情况下趋向于正态分布，这是对微积分极限过程的一次出色运用。阿兰·图灵并不知道，他的发现早已作为中心极限定理而广为人知了。不

过，他的成就给人留下了深刻的印象，这使他得到了一个研究员职位。现在，图灵成了剑桥大学的一个指导教师，年薪300英镑，任期为3年，而且几乎可以肯定能够再任3年。研究员的薪金不必与任何义务挂钩，现在图灵用餐时可以坐在“贵宾席”居高临下地面对本科生了。如果愿意，他还可以通过辅导本科生挣一些外快。这一任命把图灵引向了一条通常被认为通往学术生涯的道路。[33] 在舍波恩，当人们欢庆他们培养出来的这个孩子的成功时，所谓数学的难闻气味以及不要只成为一个科学专家的警告被忘得一干二净。学生们可以放半天的假，以下无耻的诗句也随之不胫而走：

> 图灵一定打早就在诱骗，
>
> 为的是能当上一个老师。[7]

没过多久，阿兰·图灵的确用他所掌握的那些全新的数学证明了他的能力，其结果是一篇发表的论文。巧的是，在这篇论文中，他对冯·诺依曼所证明的一条定理进行了改进，它属于一个极其专业的领域——殆周期函数论。图灵现在完全走在一条通往成功的数学研究人员的道路上，他的成就只会让其他专家感兴趣。1935年春，他在剑桥听了一门关于数学基础的课程，接着，一切都改变了。

希尔伯特的判定问题

莱布尼茨曾经梦想能够把人的理性还原为演算，并且有强大的机器能够执行这些演算。弗雷格第一次给出了一个似乎能够解释人的一切演绎推理的规则系统。哥德尔在1930年的博士论文中曾经证明了弗雷格的规则是完备的，这样就回答了希尔伯特2年前提出的一个问题。希尔伯特也试图找到清楚明白的演算程序，只要用所谓一阶逻辑的符号系统写出的某些前提和所提出的结论给定，那么通过这些程序就总是可以判定，弗雷格的规则是否足以保证该结论能够从这些前提中导出。[8] 寻找这些程序的任务后来被称为希尔伯特的“判定问题”。当然，用于解决特定问题的演算程序系统并不是什么新的东西。事实上，传统数学课程在很大程度上就是由这些或被称为算法的演算程序构成的。我们先是学习数的加减乘除算法，然后又学习对代数表达式进行操作和解方程的算法。如果我们继续学习了微积分，那么我们就学会了使用莱布尼茨最早

为这门学科发展出来的算法。然而，希尔伯特所要找的是一种广度前所未有的算法。从原则上讲，解决他的判定问题的算法将把人的一切演绎推理都还原为冷冰冰的演算，它在很大程度上将实现莱布尼茨的梦想。

数学家们经常喜欢从两个方向解决一个困难的问题。一方面，他们试图考虑一般问题的特殊情形，而另一方面则是把一般问题还原为某些特殊情形。如果一切顺利，这两条路将殊途同归，从而为一般问题提供一种解答。关于判定问题的工作就是沿着这两个方向进行的，而且已经找到算法的特殊情形与一般问题被还原成的特殊情形之间的鸿沟已经被大大缩小，以至于再往前走一小步就可望完全填平这个鸿沟了，从而可以给出希尔伯特所要寻求的算法。[9] 剑桥的G.H.哈代对此持怀疑态度，他不无气恼地评论道：“当然不存在这样的定理，这是非常幸运的，因为它如果存在，那么我们就可以用一套机械的规则来解决一切数学问题，我们数学家的活动也就将寿终正寝了。”[10] 哈代当然不是第一个确信他的技巧永远也不会被机器替代的行家里手，但后来的结果表明，这位行家说对了！

比图灵大15岁的剑桥的另一位教师M.H.A.纽曼也是圣约翰学院的研究员，他将在图灵的职业生涯中一直扮演重要的角色。纽曼已经在拓扑学方面做出了开创性的贡献，当时拓扑学还是一门相对较新的数学分支。粗略地说，拓扑学研究的是几何形体经过任意拉伸（只要没有撕破）而保持不变的性质。纽曼在剑桥开设的拓扑学课程把许多年轻的数学家引入了这个正在迅速成长的领域，他还为此写了一本出色的教科书。当纽曼在波伦亚参加1928年的国际数学家大会时，他听到希尔伯特定下了那些目标。仅仅2年之后，年轻的库尔特·哥德尔就表明它们是站不住脚的。纽曼显然被这些进展吸引住了，他在1935年春季开设了一门关于数学基础的课程，其中最令人感兴趣的部分是哥德尔的不完备性定理。在听这门课的过程中，图灵得知了希尔伯特的判定问题。撇开像哈代那样的怀疑不谈，在哥德尔的工作发表之后，人们很难相信希尔伯特所希望的那样一种算法是存在的。阿兰·图灵开始思考如何能够证明这样的算法是不存在的。

图灵对计算过程的分析

图灵知道，一种算法往往是通过一系列规则说明的，人们可以以一

种精确的机械方式遵循这些规则，就像按照菜谱做菜一样。但图灵把关注的焦点从这些规则转移到了人在执行它们时实际所做的事情。他能够说明，通过丢掉非本质的细节，人们可以局限在少数几种极为简单的基本操作上，而不会改变最终的计算结果。接下来，图灵要说明这个人可以被一个能够执行这些基本操作的机器所替代。然后，只要证明仅仅执行那些基本操作的机器不可能判定一个给定的结论是否可以用弗雷格的规则从给定的前提中导出，他就能够下结论说，判定问题的算法是不存在的。作为副产品，他发现了通用计算机器的一个数学模型。

为了尽可能地弄清楚图灵的思考过程，让我们想象自己正在观察一个计算过程。正在做计算的人实际上在做什么呢？她（因为20世纪30年代似乎做这些工作的都是女性）正在一张纸上做记号。[34] 我们会看到，她正在把注意力从以前写的东西转移到当前正在写的东西。图灵试图把不相干的细节从这种描述中去除。她在工作时是否在喝咖啡？这当然是不相干的。她是用铅笔还是用钢笔写字？这也无关紧要。那么，纸张的大小如何？如果纸张较小，那么她就很可能需要经常看看前面那些纸。但图灵确信，这只关乎方便而绝非必需。即使她所用的纸张小到无法在一个符号下面写下另一个符号，或者说，即使她用的是一卷被分成水平方格的纸带，情况也不会发生任何本质的变化。为简便起见，我们设想她正在计算一个乘法：

$$\begin{array}{r} 4231 \\ \times 77 \\ \hline 29617 \\ 296170 \\ \hline 325787 \end{array}$$

我们可以想象她正沿着一个被划分成方格的纸带进行工作，这不会丢失任何本质性的东西，比如

4	2	3	1	×	7	7	=	2	9	6	1	7	+	2	9	6	1	7	0	=	3	2	5	7	8	7

图灵确信，虽然沿着这样一个一维纸带进行一个复杂的演算也许是很麻烦的，但这样做并不会导致什么基本问题。我们继续来观察计算过程，现在把注意力集中在一卷纸带：我们的目光沿着纸带前后移动，写下符号，有时还会返回去擦除符号并在原处写下新的符号。她关于下一步要写什么的决定将不仅取决于她正在注意的符号，而且还取决于她当前的心灵状态。即使是在我们这个简单乘法的例子中，当她注意到一对对的数字时，她的心灵状态将决定是把它们相加还是相乘。开始时，纸带看起来是这样的：

			⇓			⇓	
4	2	3	1	×	7	7	=

数字1和7上方的箭头（⇓）表示她最初注意的是这些符号。她在纸带上写下它们的乘积7：

		⇓				⇓		
4	2	3	1	×	7	7	=	7

现在她把注意力转向数字3和7，接下来轮到它们相乘了。像这样把一对对的数字相乘之后，她需要把得到的两部分乘积加起来：

												⇓							⇓	
4	2	3	1	×	7	7	=	2	9	6	1	7	+	2	9	6	1	7	0	=

她先把数字7和0加起来，得到

											⇓							⇓			
4	2	3	1	×	7	7	=	2	9	6	1	7	+	2	9	6	1	7	0	=	7

现在她必须把数字1和7加起来得到8。注意，这时她所注意到的数字与她开始演算时进行相乘的数字是相同的。然而尽管数字是相同的，但她的心灵状态却不一样，这使她去做加法而不是乘法。

这个简单的例子已经说明了任何计算所具有的本质特征。每一个进行算术、代数、微积分或任何一个数学分支中的计算的人的操作都要受

到以下限制：

·在计算的每一个阶段，只有少数符号受到了注意。

·每一个阶段所采取的行动仅仅取决于受到注意的那些符号以及计算者当前的心灵状态。

一个人可以同时处理多少符号？正确地进行一项计算需要多少人？第一个问题的答案当然是因人而异的，但无论如何不可能很多。而对于第二个问题，答案是一个人。这是因为，同时注意几个符号总是可以通过每次注意一个符号来实现。[11] 而且，把注意力从纸带上的一个方格转到一定距离以外的另一个方格总是可以通过一系列移动来实现，其中每一次移动或者是左移一个方格，或者是右移一个方格。由这一分析可以推出，任何计算都可以被看作是以下面的方式进行的：

·计算通过在一条被划分成方格的纸带上写下符号来执行。

·执行计算的人在每一步都只注意其中一个方格中的符号。

·她的下一步将仅仅取决于这个符号和她的心灵状态。

·她的下一步是这样的：她在当前注意的方格里写下一个符号，然后把注意力转向它左边或右边的相邻方格。

现在就很容易看出，做这项工作的人可以用一个机器来替代，纸带（可以看成带有用编码信息来表示的符号的磁带）在机器中来回移动。计算者的心灵状态用机器内部的不同配置来表示。机器的设计必须使得每一时刻纸带上只有一个符号被感知，这个符号被称为被注视符号。根据其内部配置和被注视符号，机器将在纸带上写下一个符号（取代被注视符号），然后或者继续注视同一个方格，或者转到纸带上与之相邻的左边或右边的位置。就计算的目的而言，机器如何制造或者由什么来制造，这都无关紧要，重要的只是它能够具有不同的配置（也被称作状态），当它处于每一种的配置或状态时，都可以做恰当的处理。

问题的关键不在于真的造出一台图灵机，它毕竟只是数学的抽象。[35] 关键之处在于，基于图灵对计算概念的分析，我们就有可能得出结论说，通过任何算法程序可计算的任何东西都可以通过一台图灵机来计

算。因此，如果我们可以证明某项任务无法用图灵机来完成，那么我们就可以说，没有任何算法程序可以完成这项任务。这就是图灵证明判定问题不存在算法的方法。此外，图灵还说明了如何造出这样一台图灵机，这台图灵机可以独自完成任何图灵机所能做到的任何事情——通用计算机的一个数学模型。

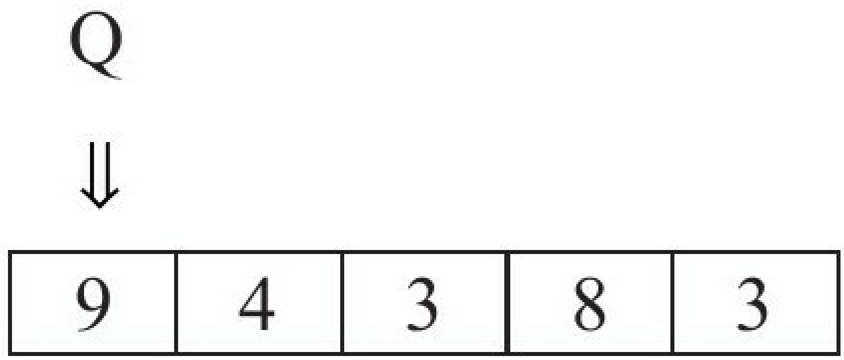
运转的图灵机

由图灵对计算过程的分析我们可以看出，任何计算都可以通过一种后来被称为图灵机的受到严格限定的装置来实现。我们可以考察几个非常简单的例子。演示一台图灵机需要哪些东西呢？首先，我们需要它所有可能的状态。然后，对于每一个状态以及纸带上可能出现的每一个符号，我们必须指明处于那个状态的机器遇到那个符号时会怎样进行操作。我们已经说过，这个操作仅仅包括被注视方格中符号的可能改变、向左或向右移动一个方格以及状态的一种可能变化。如果我们用大写字母来表示不同的机器状态，那么陈述：

> 当机器正处于状态R，注视纸带上的符号a时，它将用b来代替a，向右移动一个方格，然后转到状态S。

就可以用公式表达为Ra：b→S。类似地，同样条件下向左移动一个方格可以表示为Ra：b←S。最后，只改变纸带上的符号而不沿纸带运动可以表示为Ra：b*S。我们通常把这些公式称为“五元组”，因为其中的每一个都要用到5个符号（不包括冒号）。于是任何一台图灵机都可以通过一系列这样的五元组来表示。

让我们看看如何造出一台能够检验一个给定的自然数是奇数还是偶数的图灵机。所指定的数将用我们所熟悉的记法，记作由1，2，3，4，5，6，7，8，9，0所组成的一串数字。当然，（以这种方式写出的）一个数是奇是偶一眼就能看出。只要看看最右边的数字，如果它是1，3，5，7或9，那么这个数就是奇数，否则就是偶数。但是我们所使用的装备将从注视最左边的数字开始。由于图灵机每次只能处理一个数字，而且每次只能移动一个方格，所以如何进行处理并不是显而易见的。写在纸带上的“输入的”数就像这样：

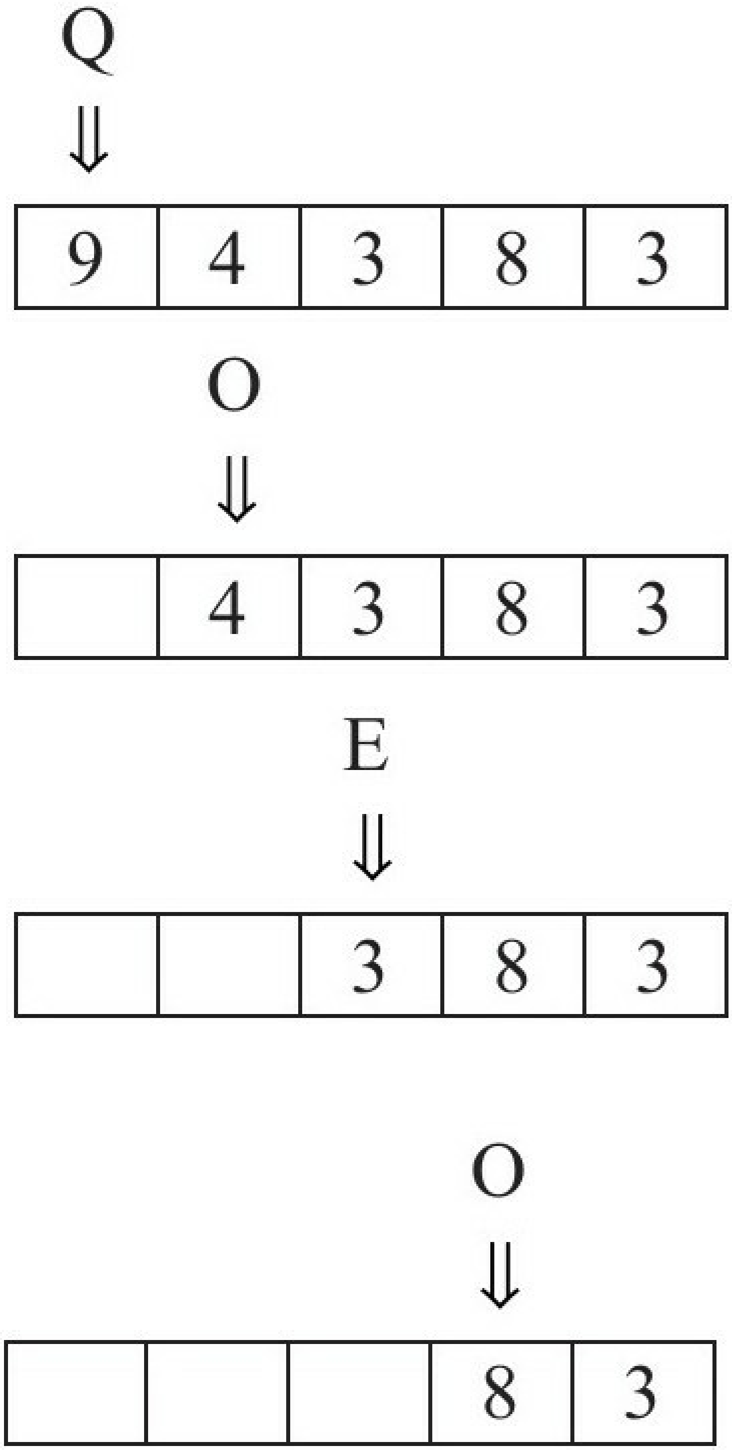


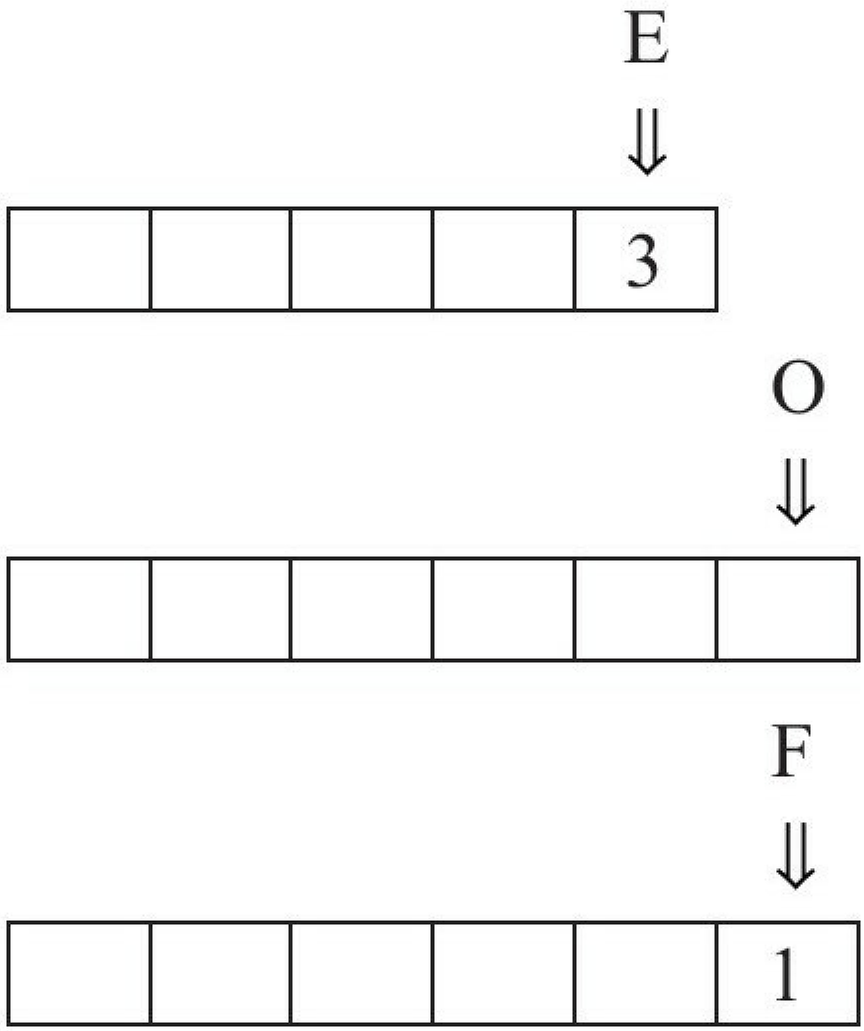
这里纸带上写的是数字94383，机器处于它初始的状态Q上，此时正在注视最左边的方格。虽然纸带只显示了5个方格（刚好足够容纳输入的数），但对于一项计算来说，很重要的一点就是纸带的方格数量没有限制。因此，如果机器试图从纸带的最右端移出去，那么一个空白的方格总会出现。我们把空白方格当作一个特殊字符来处理，记作□。

我们的图灵机将总是从注视最左端方格的状态Q开始。无论什么数输入机器，最后的结果都是一条除一个方格外其余都是空白的纸带。如果最初输入的数是偶数，那么那个方格中将是0；如果最初输入的数是奇数，那么那个方格中将是1。这台机器将具有4个状态，分别用Q，E，O和F来表示。我们已经说过，Q是初始状态。无论机器处于什么状态，如果它注视的是一个偶数，那么它将擦除这个数字（即在它上面印上一个空格），向右移动一格，然后终止于状态E。类似地，如果它注视的是一个奇数，那么它将擦除这个数字，向右移动一格，然后终止于状态O。最终，它将注视并擦除整个输入，得到一个空格。如果它终止于状态E，那么它将印上一个0；如果终止于状态O，则将印上一个1。然后，它将向左移动一格停止下来。下面是组成这个机器的五元组：

Q0：□→E	Q2：□→E	Q4：□→E	Q6：□→E	Q8：□→E
Q1：□→O	Q3：□→O	Q5：□→O	Q7：□→O	Q9：□→O
E0：□→E	E2：□→E	E4：□→E	E6：□→E	E8：□→E
E1：□→O	E3：□→O	E5：□→O	E7：□→O	E9：□→O
O0：□→E	O2：□→E	O4：□→E	O6：□→E	O8：□→E
O1：□→O	O3：□→O	O5：□→O	O7：□→O	O9：□→O
E□：0*F	O□：1*F			

输入为94383的全部计算显示了机器的详细操作：

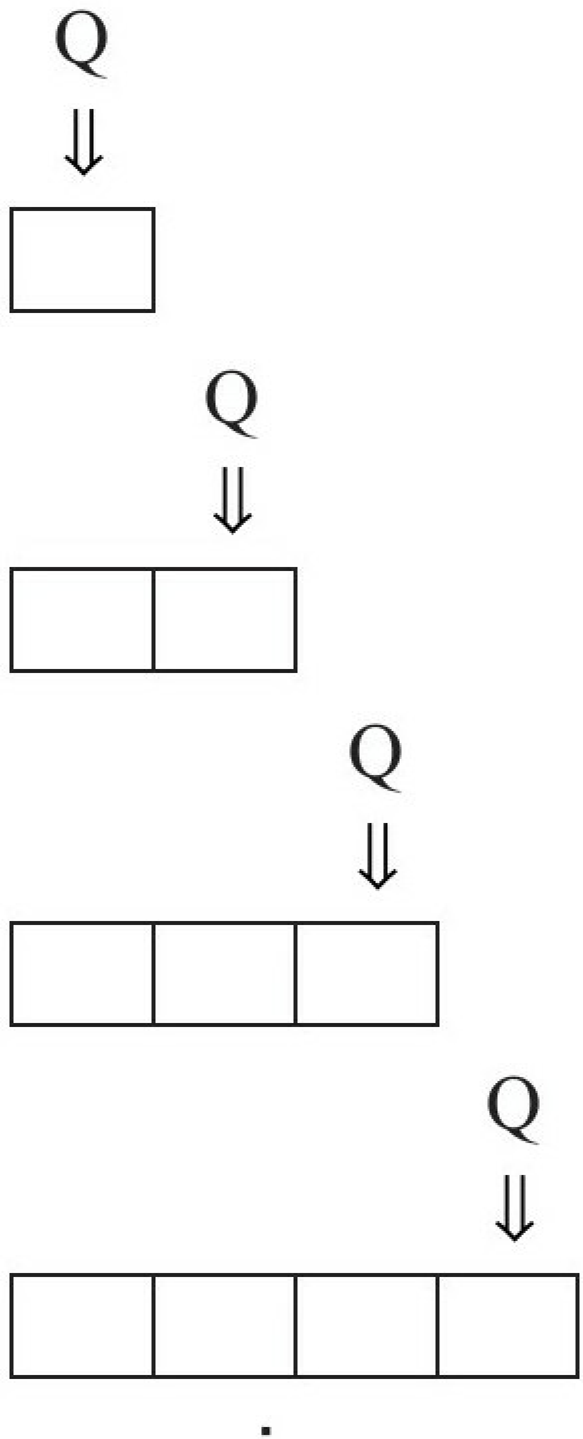

Q
9 4 3 8 3
O
4 3 8 3
E
3 8 3
O
8 3



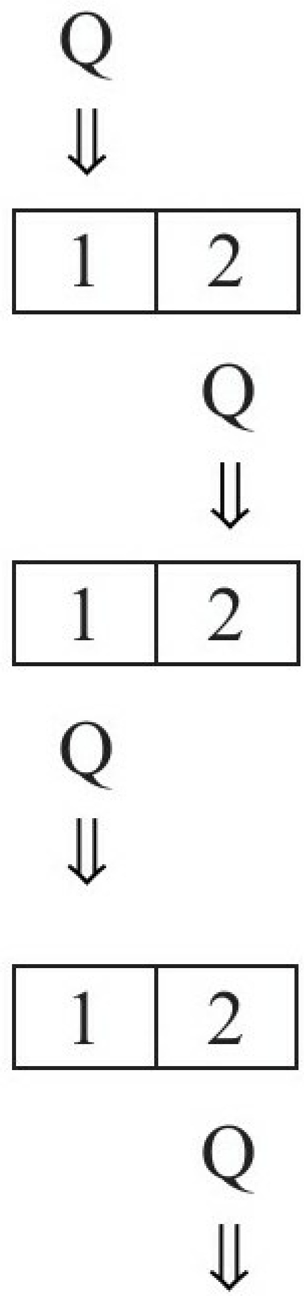
机器从注视数字9的状态Q开始，相应的五元组位于表中的第二行、最后一列。这一五元组使机器擦除了9，向右移动一个位置，进入状态O。在注视4的状态O中，相应的五元组位于表中的第五行、第三列。相应地，机器擦除4，向右移动，进入状态E。接着，在注视3的状态E中，相应的五元组位于表中的第四行、第二列，机器擦除3，继续向右移动，进入状态O。在注视8的状态O中，相应的五元组位于表中的第五行、最后一列，机器擦除8，向右移动，进入状态E。状态E又一次注视3，相应的五元组位于表中的第四行、第二列，机器擦除3，继续向右移动，进入状态O。在状态O中，机器注视的是一个空格，相应的五元组位于表中的最后一行、第二列，空格被1替换，机器处于原位不动，进入状态F。在状态F中，机器注视的是一个空格，此时没有五元组可用

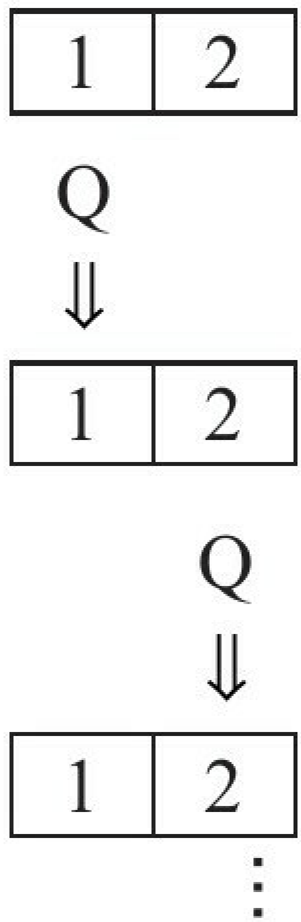
了，于是机器停止。计算的结果是纸带上只留有一个数字1，这是正确的，因为输入的数为奇数。

与物理装置不同，图灵机仅仅作为数学抽象而存在，所以可以不必受纸带的量的限制。当由一个五元组Q□：□→Q所组成的图灵机从一条空白的纸带开始时，它将随着纸带的量的持续增长而“永远”向右移动：

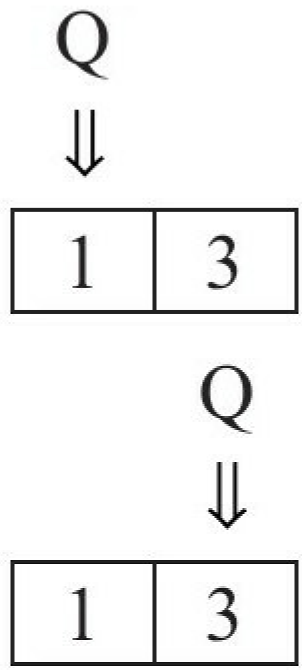


图灵机的计算可以永远持续下去，即使它所用的纸带的方格数是固定的。例如，考虑由Q1：1→Q和Q2：2←Q这两个五元组所组成的图灵机。如果输入为12，那么这台机器将像下面这样来回跳动：


Q
1
2
Q
1
2
Q
1
2
Q



这种行为严格依赖于输入。例如，如果输入是13，那么同一台机器的计算将是下面这样：



在注视3的状态Q中，没有五元组可以应用，于是机器停止。

总之，某些输入会使有些图灵机停止下来，另一些则不会。图灵把康托尔的对角线方法应用于这种情况，得到了图灵机所不能解决的问题，由此便推出了判定问题的不可解性。

图灵应用康托尔的对角线方法

马科斯·纽曼的课使阿兰·图灵注意到了判定问题，它最令人感兴趣的部分是哥德尔的不完备定理。所以正在思考用一系列五元组来表示他的机器的图灵就很自然地想到了用自然数来做机器的代码，并且可以使用康托尔的对角线方法。我们将遵循图灵的思路，设置一种与他所使用的代码相似但不尽相同的代码。

为了设置我们的编码方案，我们考虑彼此由分号隔开的五元组系列来建立图灵机。于是由一对五元组：

Q1：1→Q Q2：2←Q

所组成的图灵机可以记作Q1：1→Q；Q2：2←Q。然后我们根据下述方案用一串十进制数来替换每一个符号：

·开始和结尾都为8，其间只有0、1、2、3、4、5的字串将被用来表示纸带上的符号。下表给出了我们用来表示十进制数字和□（作为纸带上的符号）以及五个符号→←*：；的特殊表示：

符号	表示	符号	表示
0	8008	□	8558
1	8018	→	616
2	8028	←	626
3	8038	*	636
4	8048	:	646
5	8058	;	77
6	8518		
7	8528		
8	8538		
9	8548		

·开始和结尾都为9，其间只有0、1、2、3、4、5的字串将被用来表示状态。特别地，开始的状态Q将用字串99来表示。

于是，刚才所说的两个五元组的图灵机将用998018 646 8018 616 99 77 998028 646 8028 626 99来编码。对于我们用来区分奇数和偶数的图灵机来说，我们可以分别用919、929和939来编码E、O、F这三个状态。这样，整个机器的码数就成为：

998008646855861691977 998028646855861691977 998048646855861691977

998518646855861691977 998538646855861691977

9980186468558616929 77

9980386468558616929 77 9980586468558616929 77 9985286468558616929 77

9985486468558616929 77 91980086468558616919 77 91980286468558616919 77

91980486468558616919 77 91985186468558616919 77 91985386468558616919 77

91980186468558616929 77 91980386468558616929 77 91980586468558616929 77

91985286468558616929 77 91985486468558616929 77 92980086468558616919 77

92980286468558616919 77 92980486468558616919 77 92985186468558616919 77

92985386468558616919 77 92980186468558616929 77 92980386468558616929 77

92980586468558616929 77 92985286468558616929 77 92985486468558616929 77

91985586468008636939 77 92985586468018636939

尽管这不过是一个大数，但它已经用空格显示了各个五元组的代码数。请注意，从这个代码数中恢复五元组是很简单的：首先找到分隔五元组码数的77这个字串，然后再对每一个五元组进行解码。例如：92985386468558616919这个码数可以分解成929 8538 646 8558 616 919，它可以被解码为O8：□→□E。当然，编码可以通过许多不同的方式来设置，但这一方案拥有一个重要而有用的性质，那就是明显的可解码性。[36]

从上面的例子可以看出，任何图灵机都可以被认为是从纸带上的数

的最左边的数字开始注视。在这些数中，有些数会使机器最终停止，而另一些则会使机器永远运行下去。让我们把由前一类自然数所组成的集合称为特定图灵机的停机集合。现在，如果我们认为一台图灵机的停机集合组成了一个“包裹”，并认为那台机器的代码数就是这个包裹的标签，那么我们就具备了应用对角线方法的标准条件：包裹上的标签就是包裹里的东西——这里是自然数。[37] 对角线方法将允许我们构造出一个与图灵机的任何停机集合都不同的自然数集合，我们把它称为D。方法是这样的：D将完全由图灵机的代码数组成。对于每一台图灵机来说，它的代码数属于D，当且仅当它不属于那台机器的停机集合。这样，如果某台图灵机的代码数属于它的停机集合，那么这个代码数就不属于D；而如果这个代码数不属于机器的停机集合，那么它就属于D。无论是哪种情况，D都不可能是这台机器的停机集合。既然每一台图灵机都是如此，我们就可以得出结论说，集合D不是任何图灵机的停机集合。

等等！这里还有一个固执的人不相信这一点。我们听听在这个固执的人（SP）和无所不知的创造者（OA）之间展开的一场对话：

SP：我没有完全听懂这一推理，但无论如何，我知道我可以构造出一台停机集合为D的图灵机。这就是它。

OA：那么，你是否可以演算一下你的机器的代码数？

SP：好啊！让我想一想。这个数是99803864685586192977……7792985286468558616929（给我们显示出某个巨大的数）。

OA：好的。这个数在你的机器的停机集合当中吗？

SP：等等！我必须把它算完。不，不。它不在我的机器的停机集合当中。

OA：且听我说。如果这个数不在你的机器的停机集合当中，那么根据D的定义，这个数就必然在D当中。由于这个数在D当中，而不在你的机器的停机集合当中，所以这两个集合必定是不同的。

SP：让我检查一下。哦，我明白了，我犯了一个小错误，真傻。这个数其实就在我的机器的停机集合当中。我要为我愚蠢的错误道歉。

OA：别急！根据D的定义方法，如果你的机器的代码数在它的停机集合当中，那么它就必然不在D当中。因此这两个集合必定是不同的。

SP：你说的真是很有道理。但如果我同意你已经证明了你的观点，那么我就不再是一个顽固的人了。

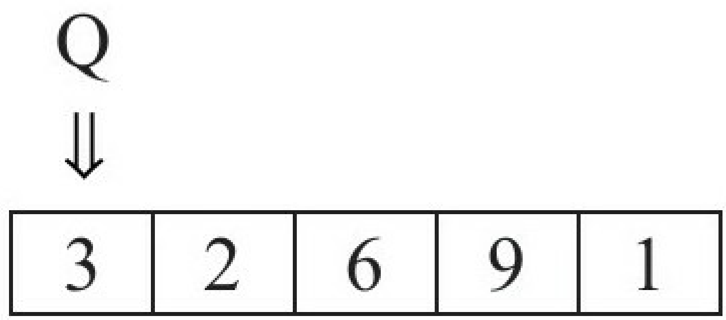
不可解问题

自然数集D的定义使它与任何图灵机的停机集合都不同。但这与判定问题有什么关系呢？它们之所以有关，是因为希尔伯特把这个问题称为数理逻辑的基本问题。希尔伯特认为，对判定问题的解答将会为解决所有数学问题提供一种算法。同样的理解也潜藏于哈代的看法中，他确信判定问题永远也不会有一个解答。只要我们严肃地看待这一点，便会发现，它隐含着只要有一个数学问题可以被证明在算法上是不可解的，那么判定问题本身就必定是不可解的。集合D将为我们提供这样一个例子。

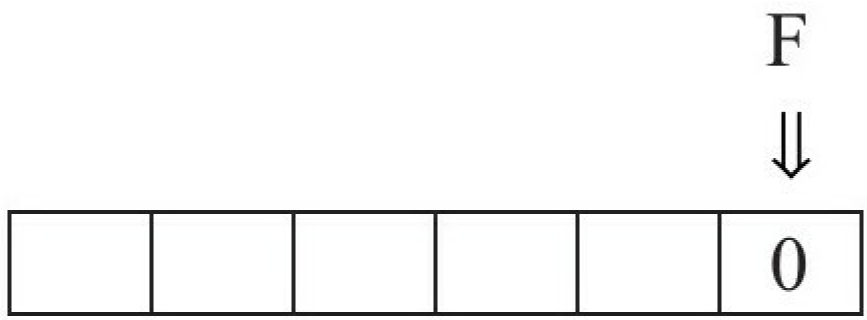
考虑下面这个问题：

找到一种算法，判定一个给定的自然数是否属于集合D。

这就是一个不可解问题的例子。要想弄明白不存在这样的算法，我们首先要看到，通过图灵对计算过程的分析，如果存在着这样一种算法，那么就会有一台图灵机能够完成同样的事情。就像我们构造出来的区分奇偶数的图灵机那样，我们可以想象这样一台处于初始状态Q的机器从已知数的最左端数字开始注视，就像这样：



类似地，我们将希望机器最终会停机，除了有一个数字，纸带的其余部分都是空白的：如果输入的数属于集合D，那么这个数字就是1，否则就是0。最后，我们希望它终止于状态F，并且机器的五元组都不从字母F开始。[38] 例如，



现在，我们想象把以下两个五元组加到我们的图灵机中：

F0：□→F和F□：□→F

如果输入的数属于D，那么新机器将像以前那样运转，最终将以纸带上的1而告终。然而，如果输入的数不属于D，那么这台机器将永远向右移动。因此，这台新机器的停机集合恰恰就是集合D。然而这是不可能的，因为D是用对角线方法构造出来的，所以它与任何图灵机的停机集合都不同。因此，我们关于存在着一种区分D的成员与非成员的算法的假定必定是错误的。这样的算法不存在！用算法来区分D的成员与非成员的问题是无法解决的！

正如我们已经看到的，希尔伯特和哈代都相信，对于判定问题的一个算法解将意味着任何数学问题都可以通过一种算法来解决。所以一旦我们有了一个在算法上不可解的数学问题，那么判定问题的不可解性就得到了。为了看出如何与集合D发生关系，我们把下面的前提和结论与每一个自然数n相联系：

<u>前提</u>

n是某台图灵机的代码数，它被印在机器的纸带上，最左边的数字

被注视。

结论

以这种方式启动的图灵机最终将停止。

利用一阶逻辑的语言，这两个句子都可以被翻译成逻辑符号。于是可以证明，这个结论可以用弗雷格的规则从前提中导出，当且仅当这台图灵机从它纸带上的代码数开始后将最终停止。相应地，它为真当且仅当n不属于D。因此，如果我们有了一种判定问题的算法，那么我们就可以用它来判定一个数是否是D的成员。也就是说，给定一个自然数n，我们可以用这种判定问题的算法来检验这个结论是否可以从前提中导出。如果可以，那么我们就知道n不属于D，如果不可以，那么n就属于D。由此可知，判定问题在算法上是不可解的。[12]

图灵的通用机

图灵的工作中有一些地方是令人困惑的。他已经证明，没有图灵机可以被用于解决判定问题。然而，为了说明判定问题不存在任何种类的算法，图灵曾诉诸他对人执行一项演算时的过程的讨论。他关于任何这样的计算也可以用一台图灵机来完成的论证是否令人信服呢？为了支持他的论证，图灵证明了许许多多复杂的数学演算都可以在图灵机上完成。[39] 不过，他在检验自己的工作的有效性方面所提出的最为大胆的、影响最为深远的思想却是通用机。

考虑被一个空白方格所隔开的写在图灵机纸带上的两个自然数（通常的十进制记法）。第一个数是某台图灵机的代码数，第二个数是这台机器的一个输入：

图灵机$\mathscr{M}$的码数 $\mathscr{M}$的输入

现在想象一个人要完成这样一项任务，他要完成把纸带上的第二个数作为输入时，代码数为第一个数的图灵机所要做的事情。这项任务是很简单的。他先得到代码数为纸带上第一个数的机器的五元组，然后只要按照五元组的要求在纸带上进行操作就可以了。图灵的分析已经证明，任何计算任务都可以通过一台图灵机来实现。把这一思想应用到当前的任务中，我们可以设想一台图灵机，它从图灵机的代码数开始，只要它的输入与的输入相同，它就可以完成机器所能做到的事情。一台图灵机单凭自身就可以完成任何图灵机可能做到的任何事情。图灵通过说明一个人如何能够得出这样一台通用机的五元组来检验这个非凡的结论。在现在被称为程序设计的几页文献中，他出色地做到了这一点！[13]

自莱布尼茨的时代起，甚至更早一些时候，人们一直在对演算机器进行思考。在图灵之前，一般的想法是，对于这种机器来说，机器、程序和数据这三种范畴是完全分离的。机器是一种物理对象，今天我们把它称为硬件；程序是做计算的方案，也许体现在穿孔卡片或线路连接板上的缆线连接上；而数据则是数值输入。图灵的通用机表明，这三种范畴的相互分离是一种错觉。图灵机开始被看成是一台拥有机械部件——硬件——的机器。它在通用机纸带上的代码数则有程序的功能，它为通用机详细指明了执行适当的计算所需要的指令。最后，通用机在一步步的运转中把机器代码的数字仅仅看成需要进一步处理的数据。这三个概念之间的转换对于现在的计算机实践来说是非常基本的。用一种现代的编程语言所写成的程序对于处理它以使其指令能够得到执行的解释程序或汇编程序来说就是数据。事实上，图灵的通用机本身就可以被看成一个解释程序，因为它是通过解释一连串五元组来执行它们所标明的任务的。

图灵的分析为理解古代的计算技术提供了一种独到而深刻的洞见。计算的概念原来远不止算术和代数运算。同时，这种眼光预见到了原则上能够计算任何可计算的东西的通用机。图灵关于专用机器的例子已经成了程序设计的实例；通用机器则特别是解释程序的第一个例子。通用机还为存储程序计算机提供了一个模型，在其中纸带上编了代码的五元组扮演了存储程序的角色，机器在程序和数据之间没有做基本的区分。最后，通用机表明了如何能够把被认为是机器功能描述的五元组形式的硬件，替换成以编码的形式存储在通用机纸带上的以同样的五元组形式表现出来的等价的软件。

在证明判定问题不存在算法解的过程中，图灵没有想到在大西洋的对岸也有人得出了类似的结论。当录有普林斯顿大学的阿隆佐·丘奇 [40] 的一篇题为“一个不可解的初等数论问题”的文章的《美国数学杂志》（*American Journal of Mathematics* ）送到剑桥时，纽曼已经收到了图灵论文的第一稿。在这篇论文中，丘奇说明了在算法上不可解的问题是存在的。他的论文没有提到机器，但他的确点出了两个概念，其中每一个都是为了解释可计算性或丘奇所谓的“能行的可演算性”的直观概念而提出来的。这两个概念分别是丘奇和他的学生斯蒂芬·克林所提出的λ可定义性，以及哥德尔（在他1934年春访问普林斯顿高等研究院期间所做的讲演中）所引入的一般递归性。这两个概念已经被证明是等价的，丘奇的不可解问题实际上是相对于其中的某一个概念不可解。尽管丘奇在这篇论文中并没有得出希尔伯特的判定问题本身相对于这些概念都是不可解的结论，但《符号逻辑杂志》1936年的第一期中刊出了丘奇的一个简注，在这个注释中，他的确得出了这个结论。图灵很快就证明了他的可计算性概念与λ可定义性是等价的，并且决定到普林斯顿去待一段时间。

虽然图灵的大部分成就都可以说是对美国人所做工作的再发现，但他对计算概念的分析以及他对通用计算机的发现却是全新的。 [41] 库尔特·哥德尔曾经对丘奇的理论颇有疑虑，正是图灵的分析才最终使他相信它们是正确的。 [14]

阿兰·图灵在普林斯顿

虽然英国的数学家们通常并不在乎能够获得博士学位，但图灵去普林斯顿大学的最方便的办法就是在那里当一名研究生，这与他的成就真是不相配。他在普林斯顿的2年里完成了一篇著名的博士论文（阿隆佐·丘奇任导师）。由于从系统外部考察时，哥德尔的在一给定系统中的不可判定命题可以被看成为真，所以一个自然的想法就是把这样一个命题作为一个新的公理加到已知系统中，这样就得到了一个新系统，其中不可判定命题不再是不可判定的。当然，应用哥德尔的方法，新系统也会有它本身的不可判定命题。图灵在他的学位论文中一次次地这样做，从而研究了系统的分层结构。

这篇学位论文所引入的另外一个概念是一种改进了的图灵机，它可

以中断它的计算来寻求外部信息。利用这样的机器，谈论两个不可解问题中哪一个“更不可解”就是可能的了。总而言之，这篇论文所引入的思想将会为一系列研究者的工作提供基础。[15]

1936年的时候（一直到20世纪50年代），普林斯顿大学数学系坐落在一座漂亮的红砖建筑范氏大楼（Fine Hall）中。[42] 那时，范氏大楼里不仅有普林斯顿大学的数学教师，而且还有新近成立的高等研究院的数学家。从纳粹统治之下逃出来的科学家已经开始大量涌入美国了。20世纪30年代，数学精英齐聚普林斯顿使得哥廷根最终被超越。在范氏大楼的走廊上我们能够见到赫尔曼·外尔、阿尔伯特·爱因斯坦和约翰·冯·诺依曼，后者的兴趣已经偏离希尔伯特关于数学基础的纲领很远了。

在普林斯顿的第一年，剑桥提供给他的奖学金很少，图灵必须想办法应对。当他在剑桥时，奖学金还是相当丰厚的，而且学校还提供食宿。不过在第二年，他觉得自己已经很富裕了，因为他获得了享有盛誉的普罗克特（Proctor）奖学金。在支持他申请这一奖学金的推荐信中有这么一封：

1937年6月1日

先生，

A.M.图灵先生告诉我，他正在申请1937—1938年度的剑桥赴普林斯顿大学的普罗克特访问奖学金。我很愿意支持他的申请，我认识图灵先生已经有些年头了：我在剑桥做访问教授的1935年的第二个学期，以及图灵先生1936—1937年在普林斯顿逗留期间，我都有机会看到了他的科学工作。他在我所感兴趣的数学分支即殆周期函数论和连续群理论方面工作出色。

我认为他是普罗克特奖学金的一个非常合格的候选人，如果您发现有可能授予他该奖学金，我将非常高兴。

约翰·冯·诺依曼敬上。[16]

奇怪的是，既然冯·诺依曼曾经深入研究过希尔伯特关于数学基础的纲领，图灵关于可计算性的工作以及他关于判定问题不可解性的证明竟然没有在这封信中被提及。很难相信冯·诺依曼不知道这一点。我认

为其中的关键就在于“我所感兴趣的数学分支”这几个字：作为那个世纪的一个伟大数学家，一个无书不读而且几乎过目不忘的人，在哥德尔证明了他在这个领域的大部分工作都是徒劳无益的之后，冯·诺依曼显然决定不再与逻辑打交道了。他甚至有一句名言说，在哥德尔1931年的工作出来之后，逻辑方面的文章他一篇也不读了。[17] 鉴于图灵的工作对冯·诺依曼在第二次世界大战期间和第二次世界大战之后关于计算机的思想所起的作用，这件事是很重要的。

冯·诺依曼的朋友兼合作者斯坦尼斯拉夫·乌拉姆给图灵的传记作家安德鲁·霍奇斯所写的一封信提供了一些证据。[43] 这封信提到了冯·诺依曼提出的一个游戏，这是在1938年夏天冯·诺依曼和乌拉姆一起在欧洲旅行期间提出的。这个游戏包括“在一张纸上写下尽可能大的数，并通过与图灵的某种图解有关的方法来定义它”。乌拉姆的信中还说：“……冯·诺依曼在1939年关于发展形式数学系统的机械方法的谈话中和我多次谈到了图灵。”乌拉姆的信说得很清楚，无论以前的情况怎样，到了1939年9月第二次世界大战爆发，冯·诺依曼已经非常了解图灵关于可计算性的工作。[18]

图灵的通用计算机是一个奇妙的概念装置，它仅凭自己就可以执行任何算法任务。但我们真能制造出这样一种东西吗？除了这种机器原则上所能完成的任务之外，它的设计和制造能否使其在可以接受的时间内应用适量的资源来解决真实世界的问题呢？这些问题从一开始就被图灵注意到了。在（伦敦的）《时代》杂志上刊登的一篇悼文中，图灵的老师马科斯·纽曼曾经写道：

> 他那时对“通用”计算机的描述完全是理论上的，但图灵对一切实验的强烈兴趣不禁使他好奇，沿着这些思路是否可能真的造出一台机器。[19]

他并没有一味地思考这种可能性。为了熟悉现有的技术，图灵特地用机电继电器制造出了一台能够把二进制数进行相乘的装备。为此，他进入了物理系研究生的机加工车间，制作了各种部件，而且亲手制造出了所需的继电器。[44]

阿兰·图灵的战争

1938年夏天，图灵回到了剑桥。尽管战争一年以后才会降临，但他却应邀参与了破译德军通信密码的工作。图灵和哥德尔的工作都包含了编码和解密，但那些编码故意选择得很容易识破，而不像德国人正在使用的编码努力不让人猜透。事实上，德国在整个战争中都以为他们的密码是无法识破的。

根据纳粹德国和苏联之间的一项令世界震惊的协定，1939年9月1日，德军入侵波兰。几天以后，英国和法国根据承诺向德国宣战。9月4日，图灵给布莱奇利庄园（Bletchley Park）——伦敦北部的一座维多利亚时代的庄园——做了汇报，那里有一个主要由学术界人士组成的小组，当时他们正设法读取敌军竭力向他们隐藏的信息。这个小组并不打算维持小规模，到了战争结束，大约有1.2万名从事破解密码和信息分析各个方面工作的人住在这个庄园里。其中除了军队和老员工之外，还有大量的“皇家海军女子服务队成员”，她们加入了海军后备军，操纵的正是图灵和他的同事所设计的机器。

德军用来通信的是一台被称为“谜”（Enigma）的改良的商用加密机器。这台机器有一个字母键盘，当一个表示特定字母的键被按下之后，一个小窗口里就会出现一个字母，即原始字母的加密。当整个一段话被加密之后，它将通过普通的无线电报发送出去。指定的收信人将在另一台“谜”上输入被加密的字母，于是原始信息就出现了。机器内部装有若干旋转的轮子，以便逐个字母地把输入字母变成该字母对应的加密后的字型。军用加密字型的安全保障通过外加的一块线路连接板而得到了提高。每一天机器都会有不同的初始设置，但对于发送者和接收者来说必须是相同的。

战争开始前，一群波兰的数学家已经成功地破译了关于德国的“谜”的信息。但是当德国人把机器变得更为复杂之后，他们就无计可施了，并把工作交给了英国人。布莱奇利庄园的密码专家大多是喜爱解谜的人，他们有时会全神贯注于问题的智力层面而乐在其中。但这项工作是极为严肃的，图灵的特殊责任就是在德国的潜艇和他们的总部之间传递信息。被这些潜艇击沉的给英国运送急需物资的船不计其数，如果德军潜艇不被阻止，英军就有可能因粮草不济而失利。借助于一本从被俘潜艇中发现的电码本以及发送者无意中透露出来的重要信息，“谜”的通信得以成功破译。而图灵着实功不可没，他成功地设计出一台机器（不知是什么原因，它被称为“霹雳弹”[BOMBE]），这台机器可以用这

些信息迅速地推导出“谜”在某一天的设置情况。“霹雳弹”通过系统地执行一连串逻辑推理，从大量可能的数值中把谜的可能状态一一排除，最后只剩下几个数值留给手工处理，直到最终获得正确的结果。[45]

图灵的“霹雳弹”有几个令人称奇的地方。在图灵的设计草图出来之后仅仅几个月，成打的机器就造了出来并被交付使用。令人吃惊的是，它们无须任何改进就可以工作。要想获得德国海军的“谜”在某一天的设置，就要从150 000 000 000 000 000 000种可能性中找到正确的组合。平均说来，图灵的“霹雳弹”可以在3小时之内解决这个问题，有一次它只用了14分钟。

在布莱奇利庄园，图灵被戏称为“那个教授”（the prof），他的种种怪癖也成了大家津津乐道的话题。许多年以后，人们还会谈起他把茶杯与暖气片系在一起的习惯。在布莱奇利庄园的岁月里，最富有启发性的事情也许莫过于图灵学习用步枪射击的过程了。在1940和1941年的黑暗日子里，英国似乎随时可能遭袭，丘吉尔政府组建了一个民兵组织——民团。尽管由于图灵工作的重要性，他不需要加入地方军，但他还是决定这样做，为的是学会用步枪射击。民团的新兵都要求参加按时训练，没过多久，图灵就发现这纯粹是浪费时间，于是就不参加了。当怒气冲冲的菲灵汉姆上校要求他遵守纪律时，图灵耐心地解释说，他参加民团本来就是为了学习射击的，既然他现在已经成了一个出色的射手，他就不再有任何理由参加了。上校说：“但你是否参加不是你说了算的……这是你作为一名士兵的义务……你必须遵守军纪。”上校提醒图灵，在申请入伍的时候，他曾填过一张表，表中有这样一个问题：“你知道参加地方军就意味着你必须遵守军纪吗？”图灵说，他的确回答了这个问题，但他写的答案却是：“否。”当时在考虑这个问题时，图灵显然认为回答“是”对他没有任何好处。[20]

除了有趣，这则逸事还揭示出阿兰·图灵的许多性格。他倾向于不理睬我们大部分人生活于其中的社会框架，在任何情况下，他都会通过思考得出答案，并且从头开始寻求最佳的行动。大多数人在面临类似于民团申请表上的问题时都会认为，只有一个肯定的答案才是可以被接受的，但图灵却从字面上来理解这个问题，认真思考最好的答案应当是什么。尽管对于图灵来说，这种思维方式在他的科学研究中很管用，但在他的人际交往和社会关系中却并不好使。几年以后，它最终导致了灾难。

图灵发现自己与布莱奇利庄园中的一位年轻数学家琼·克拉克变得非常友好。事实上，他爱上了她，而且提出了结婚的请求，并被愉快地接受了。几天以后，他把自己的同性恋倾向告诉了她，此时她非常不安，但还是准备履行婚约。几个月之后，就在他们一起外出度假后不久，图灵认定，尽管他真心爱着琼，但这不管用，于是还是解除了婚约。显然，这是他第一次，也是最后一次允许自己想象与一位女性有恋爱关系。

与此同时，图灵从未停止思考通用机概念的可应用性。他猜想，正是这种通用性想法揭示了人的大脑拥有强大的能力的秘密，从某种意义上来说，我们的大脑正是通用机。他设想，如果一台通用机可以被制造出来，那么它就可以玩象棋这样的游戏，可以引导它像一个孩子那样学习许多东西，并且最终展示出智能行为。在布莱奇利庄园中，人们就这些思想曾经有过许多讨论，图灵甚至还草拟了可以让机器下棋的算法。同时，制造一台通用计算机所需要的某些硬件也在布莱奇利庄园发展起来。

英国所截获的某些源自纳粹最高当局的讯息并不是“谜”加密的，也不是用普通电报传送的。英国人很快就意识到，它们具有电传打字机输出的特征。在这样一个系统中，一段文字中的每一个字母都是用一张纸带上的一排孔来表示的。与早期的摩尔斯电码不同，它不需要人工操作。德国人似乎正在使用一台能在单一的操作中加密和传送讯息的机器，接收者也要有一台能够解码的机器。在布莱奇利庄园，这个系统被称为“鱼”，图灵的老师马科斯·纽曼承担了这项破译任务。他所使用的某些方法被戏称为“图灵方式”（*turingismus*）[46]，以暗示其来源。但“图灵方式”要求对大量数据进行处理，要使破译发挥作用，这个过程必须非常迅速地完成。[21]

20世纪30年代，大多数美国人和欧洲人都已经有了收音机。在那些日子里，在晶体管发明之前，收音机里含有许多真空管（在英国称为“电子管”）。工作时，它们会像低瓦数灯泡一样发光，而且会变得很烫。就像灯泡一样，它们很快就被烧坏了，于是不得不进行更换。当收音机停止工作时，人们可以把真空管从它们的槽中拔出来，送到商店去测试。只有在更换了坏的真空管之后，收音机通常才能起死回生。美国无线电公司的真空管目录中列出了数百种型号的真空管，它们每一种都有各自的特点，是工程师和许多爱好者不可或缺的东西。1943年3月，

阿兰·图灵在访问了美国几个月后乘船回到了家。在美国，他帮助美国人制造了自己的“霹雳弹”，以监测“谜”的海军活动讯息。途经大西洋时，他研究了这份美国无线电公司目录以打发时间，因为人们已经发现，真空管可以实现以前由电子继电器来完成的逻辑运算。这些真空管很快：它们的电子以接近光速的速度行进，而继电器则依赖于机械运动。事实上，真空管电路已经被实际应用于电话交换，图灵与这项研究的先锋——极富才华的工程师T.弗劳尔斯进行了联系。在弗劳尔斯和纽曼的领导下，一台体现了“图灵方式”的机器很快就被做出来了。这台被称为“巨人”的机器是一项工程奇迹，它包含1500根真空管。世界上第一台电子自动演算装置诞生了。毫不奇怪，从本质上讲，它所进行的计算是逻辑的而非算术的。被截获的德军讯息以打孔纸带的形式被一个速度极快的读带机输送到机器里：当纸带通过读带机时，从纸孔中透出的光束被把信号传送给“巨人”的光电管截获。纸带必须被迅速地读出，以免减慢真空管电路的操作。弗劳尔斯的骄人成绩不仅体现在短短几个月就造出了一台能够运行的机器，而且还体现在用一个包含如此大量的真空管的机器做出了有用的工作。的确，许多人曾以为，真空管的频频失败将使这一切都不可能。

到了1945年战争结束时，图灵已经掌握了真空管电子学的应用知识。他确信可以用真空管电路来制造一台通用计算机，于是便花了大量精力去思考机器的具体实现和各种应用方面的事情。现在，他所需要的只是充分的支持和有利条件，以便把这项伟大的工程付诸实施。

第8章　研制第一批通用计算机

谁发明了计算机？

现代计算机是逻辑与工程的复杂混合体，单独挑出一人作为发明者是可笑的。然而1973年，在解决一场专利纠纷（Sperry Rand和Honeywell）的过程中，有一位法官差不多就下了这样的结论。随着我们的故事从背后的逻辑思想转到现代通用计算机的实际制造，工程问题以及能够有效应对这些问题的人便走上了前台。计算史已经给出了各不相同的说法，在继续我们的故事之前，我们最好先简要回顾一下这些人物：

约瑟夫—玛丽·雅卡尔（1752—1834）雅卡尔织布机是一种可以用一堆穿孔卡片来控制编织纹样的织布机器，它首先在法国引发了纺织业的革命，最后席卷了全世界。职业织工往往不无夸张地说，这就是第一台计算机。尽管它是一项奇妙的发明，但雅卡尔织布机与其说是一台计算机，不如说是一架自动钢琴。因为它就像自动钢琴那样，可以通过在输入介质上打孔来自动控制一台机械装置。

查尔斯·巴贝奇（1791—1871）参见第7章开头。巴贝奇建议把雅卡尔那样的穿孔卡片用到他没有制造出来的分析机上。他的想法类似于雅卡尔关于织布机的想法。

爱达·洛甫莱斯（1815—1852）她的父亲拜伦爵士在她一岁之后就再没见过她。她对数学，特别是对巴贝奇的分析机有很大的热情。她翻译了一本有关分析机的法文研究报告，并且经巴贝奇的同意，在其中加入了自己的注释。她被称为世界上第一位计算机程序员，为了纪念她，有一种主要的编程语言就是以她的名字“爱达”命名的。她把分析机与雅卡尔织布机联系起来的名言常常被人引用：

> 我们也许可以非常恰当地说，分析机编织的代数模式和雅卡尔织布机编织的花和叶完全一样。[1]

克劳德·仙农（1916—2001）仙农在麻省理工学院所做的硕士论文中，说明了乔治·布尔的逻辑代数如何能被用于设计复杂的开关电路。这篇论文“有助于把数字电路设计从一门技术转变成一门科学。”[2] 他的信息论在当代通信技术中起着至关重要的作用。仙农在计算机下棋的算法方面做了先驱性的工作。他说明了如何就可以制造出一台仅有两个状态的通用图灵机。（当我1953年在贝尔实验室做暑期工作时，仙农是我的上司。）

霍华德·艾肯（1900—1973）参见第7章开头。1944年，他为哈佛大学设计的、IBM用电子继电器制造的“自动序列受控计算机”研制成功，它可以完成巴贝奇所设想的任何任务。在为物理学家和工程师的数字压缩研制了一台专用机器之后，艾肯发现，通用机很难有效地处理这种计算。

约翰·阿塔纳索夫（1903—1995）这位衣阿华州立大学的不大引人注意的物理学家（和他的助手克利福德·贝利一起）基于美国参加二战时的真空管电子学，设计并制造出了一台专用的小型计算机。尽管这台机器只能处理非常特殊的问题，但它的重要性体现在它能够证明真空管电路对于计算的重要性。[3]

约翰·莫齐利（1907—1980）莫齐利的先见之明为世界上第一台大型数字处理电子计算机ENIAC在费城宾夕法尼亚大学的摩尔电子工程学院的成功研制打下了基础。莫齐利也是一个物理学家，他曾经访问了阿塔纳索夫在埃姆斯和衣阿华的实验室，并有机会研究了在那里制造的电子计算机。

约翰·普瑞斯伯·埃克特（1909—1995）埃克特是一位卓越的电气工程师，他的工作为ENIAC的成功研制起到了关键性的作用。

赫尔曼·戈德斯坦（1913—）数学家赫尔曼·戈德斯坦于1942年入伍，他被派到美国陆军军械弹道研究实验室当一名中尉。作为ENIAC项目的军方代表，他把冯·诺依曼吸收到了摩尔学院的小组中。在后来与埃克特和莫齐利的争论中，他支持冯·诺依曼一方。战争结束后，他成了冯·诺依曼在计算方面的工作的主要合作者。他关于计算史的书[Goldstine]强调了冯·诺依曼的地位，并因此而受到指责。（1954年，我如果要使用高等研究院的计算机，就要向他提出申请。）

厄尔·R.拉尔松（1911—）他是美国地方法院法官。1973年，他裁定埃克特和莫齐利关于ENIAC的专利权是无效的。他的意见包括：

> 埃克特和莫齐利并不是最先发明自动电子数字计算机的人，他们的机器来源于约翰·文森特·阿塔纳索夫博士的机器。[4]

约翰·冯·诺依曼与摩尔学院

正如我们已经看到的，1930年在柯尼斯堡召开的数学基础研讨会上，约翰·冯·诺依曼承担了解释希尔伯特纲领的任务。正是在这次会议上，库尔特·哥德尔抛出了令人震惊的消息，说他已经证明了包含初等数论在内的数学形式系统必然是不完备的，冯·诺依曼显然是第一个领会到哥德尔工作的重要性的人。不久以后，冯·诺依曼非常激动地给哥德尔写信说："我得到了一个非常漂亮的结果。我能够说明数学的一致性是不可证的。"冯·诺依曼发现，用哥德尔的方法可以证明，希尔伯特所设想的那些系统不足以证明其自身的完备性。正如我们已经指出的，当哥德尔接到这封信时，他已经得到了同样的结论，并在回信时寄出了一份包含这一结果的摘要。

冯·诺依曼是一个既虚荣又卓越的人，他很善于通过纯粹的理智力量在一门数学学科上留下自己的名字。他曾经在算术的一致性问题上花费了不少精力，在柯尼斯堡的研讨会上，他不遗余力地支持希尔伯特纲领。在认识到哥德尔的工作的深刻内涵之后，他又进而证明了一致性的不可证性，不料哥德尔竟走在了他的前头。这就够了。尽管他对哥德尔钦佩有加（他甚至就哥德尔的工作做过讲演），但他发誓再也不跟逻辑打交道了，据说他曾夸口说，在哥德尔之后，他从未读过一篇逻辑方面的论文，逻辑已经使他蒙羞，而他又并不习惯于受辱。即便如此，事实证明，这个誓言并没有被遵守，因为冯·诺依曼对强大的计算机器的需要最终使他又回到了逻辑。

和图灵一样，冯·诺依曼在战争期间的工作需要进行大规模的计算。然而，虽然布莱奇利庄园的密码分析工作主要强调涉及符号模式的计算，但冯·诺依曼却需要进行老式的繁重的数字处理。毫不奇怪，他欣然加入了费城电子工程摩尔学院的一个令人激动的项目——研制一台强大的电子计算机ENIAC。30岁的数学家赫尔曼·戈德斯坦让冯·诺依曼

加入了ENIAC计划。正如戈德斯坦所说，1944年夏，他们两人第一次在火车站见面，在此之后，冯·诺依曼很快就加入到了费城的讨论中。

如果说拥有1500根真空管的英国“巨人”电子计算机是一件工程奇迹的话，那么有着18000根真空管的ENIAC就简直令人震惊了。当时的传统看法认为，这样的配置不可能有效地工作，因为每过几秒钟就肯定会有一根真空管坏掉。ENIAC的首席工程师约翰·普瑞斯伯·埃克特是这个项目得以成功的关键人物，他要求每一个部件的可靠性都能达到非常高的标准。真空管在尽可能保守的功率水平上工作，以使真空管的损坏率保持在每星期3根。这件庞然大物塞满了一个巨大的房间，通过把缆线连接在一块很像旧式电话交换机的线路连接板上来编写程序。ENIAC模仿了当时最成功的计算机器——微分分析器。[5] 微分分析器并不是对由数字排成的数直接进行运算的数字设备，而是用可测的物理量（如电流或电压）把数表示出来，把各个部件连接在一起以模拟所需的数学运算。这些模拟机的精确度受到了测量仪器的限制。ENIAC是一台数字设备，它是能够处理与微分分析器相同的数学问题的第一台电子机。它的设计者使它的部件发挥与微分分析器的部件类似的作用，更大的速度和精确度则依赖于真空管电子学的接受能力。[6]

当冯·诺依曼开始与摩尔学院进行讨论的时候，ENIAC的研制成功已经没有什么重要的障碍了，人们开始把注意力转向下一台计算机，它的名称暂定为EDVAC（电子离散变量自动计算机）。冯·诺依曼立即开始研究新机器的逻辑结构。戈德斯坦回忆说：

> 埃克特很高兴看到冯·诺依曼对与新思想有关的逻辑问题如此感兴趣，这些会面属于最伟大的智力活动。
>
> 这项为新机器设计逻辑方案的工作正合冯·诺依曼的心意，正是在这里，他以前关于形式逻辑的工作才起到了决定性的作用。在他出场之前，摩尔学院的小组主要关注于技术问题，而当他来了之后，他成了逻辑问题的领导者。[7]

1945年6月，冯·诺依曼提出了他著名的“关于EDVAC的报告草案”，他实际上是主张把马上就要建造的EDVAC作为图灵通用机的一个物理模型实现出来。就像抽象装置上的纸带一样，EDVAC具有存储能力——冯·诺依曼称之为“存储器”，它既存储数据，又存储代码指令。为

了实用，EDVAC有一个能够在一步之内执行每一条算术基本操作（加、减、乘、除）的算术器件，而在图灵的原始概念中，这些操作需要用像“向左移动一格”这样的原始操作来建立。EDVAC把算术运算用在了十进制数上，而EDVAC则会享受到二进制符号可能带来的简洁。ENIAC还包含一个实现逻辑控制的器件，它把需要执行的指令（每次执行一条）从存储器转移到算术器件。计算机的这种组织方式后来被称为“冯·诺依曼结构”，今天绝大多数计算机仍然是按照这种非常基本的方案进行组织的，尽管它们的部件已经与EDVAC大相径庭了。[8]

EDVAC报告从未超出草案阶段，它从几个方面来说都是不完整的。特别是，有许多地方都应加上出处。图灵的名字从未被提及，但他的影响却是有目共睹的。EDVAC应当是通用的，这一概念被提到了不止一次。与图灵类似，冯·诺依曼也猜想人脑的某些显著能力源于它拥有通用计算机的能力。在EDVAC报告中，冯·诺依曼曾多次提到人脑和他所讨论的机器之间的类比。他注意到，真空管电路可以被设计得就像我们大脑中的神经元一样以多种方式运转，他还论述了EDVAC所需的算术器件和逻辑控制器件都可以由这种电路制成，而暂时没有考虑工程技术细节。虽然这份报告几乎完全缺少出处，但它却一再提到两位麻省理工学院的研究者于1943年发表的一篇论文，这篇论文讨论的就是这样一种关于理想化的“神经元”的数学理论。其中的一位作者后来说，他们曾直接受到图灵1936年的文章的启发（就是阐述他的通用机的那一篇），事实上，这篇文章仅有的参考文献就是图灵的那篇文章。更有启发意义的是，文章的作者还不辞辛苦地证明了通用图灵机可以用他们的理想化神经元来模拟，并认为主要是由于这一事实，他们的工作才没有走错方向。[9]

埃克特和莫齐利对冯·诺依曼只在EDVAC报告上写下了自己的名字非常不满。争论的原因之一就是，EDVAC报告中到底有多少属于冯·诺依曼个人的贡献，这个问题也许永远都无法得到最终解决。尽管埃克特和莫齐利后来否认冯·诺依曼曾经贡献过许多，但在报告发布后不久，他们写道：

> 从1944年的下半年一直到现在，我们可以幸运地请教……约翰·冯·诺依曼博士。他在关于EDVAC的逻辑控制的讨论中贡献良多，制定了指令代码，并为具体问题写出了代码指令以检验这些系统。冯·诺依曼博士还写了一份初步的报告，早期讨论的多数结果

都在其中进行了概括。……在他的报告中，物理结构和物理装置……被理想化的元素代替了，以免引起工程上的问题，因为这会分散逻辑考虑的注意力。[10]

冯·诺依曼希望他所说的机器实际上已经充分接近通用机了，这一点还有其他的证据。他强调一台计算机的“逻辑控制”对于它的“尽可能地接近通用”是至关重要的。[11] 为了检验EDVAC的普适性，冯·诺依曼第一次认真写出了一个程序，它不是要应用于数字处理（这是研制机器的主要目的），而是要有效地给数据分类。这个程序的成功使他确信，“基于现在已经掌握的证据，我们可以合理地得出结论说，EDVAC已经非常接近于‘通用’机器了，目前的逻辑控制原理相当可靠”。[12]

冯·诺依曼在EDVAC报告发表的那一年里所写的文章表明，他很清楚电子计算机的设计背后的原理的逻辑基础。有一篇文章的引言是这样的：

在本文中，我们将尝试不仅从数学的观点，而且从工程师和逻辑学家（即真正适合设计科学工具的个人或一群人）的观点来探讨[大规模计算]机器。[13]

另一篇文章显然暗示了图灵的思想，尽管它强调纯逻辑的考虑是不够的：

通过形式—逻辑的方法，我们很容易理解，足以控制和执行任何操作序列的抽象代码是存在的，它们可以在机器中单个使用，其整体又可以被问题计划者所设想。从目前的观点来看，在选择一个代码时，真正决定性的考虑是一种更为实际的性质：代码所要求的装置的简洁性、它运用于实际重要问题上的清晰性以及它处理这些问题的速度。一般性地或者从第一原理讨论这些问题将使我们离题太远。[14]

我们知道，第二次世界大战之后发展出来的计算机与早期的自动计算机之间有着本质性的区别。但这种区别的本质却很少有人理解。这些战后的机器都被设计成了适用于各种目的的通用装置，只要过程中的步骤被明确指定，它们就能够执行任何符号过程。有些过程所要求的存储器可能超出了实际所能达到的程度，要么就是速度过慢而不适用，所以

这些机器只能是图灵理想化的通用机的近似。然而，它们有一个巨大的存储器（与图灵机的无限长的纸带相对应）却是至关重要的，只有这样，指令和数据才能共存于其中。指令和数据之间的这种变动的界限就意味着，我们可以设计出把其他程序当成数据的程序。早期的程序员主要就是利用这种自由度来设计出能够更改自身的程序。在人们普遍使用操作系统和程序设计语言的分层的今天，更为复杂的应用已经成为可能。对于一个操作系统而言，它所启动的程序（比如你的字处理器或电子邮件程序）就是供它操作的数据，只要每一个程序都有它自身的存储部分，而且（当处理多重任务时）能够记录所要完成的每一项任务的进程。汇编程序把用今天常用的某种语言写成的程序翻译成可以被计算机直接执行的指令：对于汇编程序而言，这些程序就是数据。

在具备了ENIAC和“巨人”的经验之后，那些对计算装置感兴趣的人所能够满足的运算速度将不会低于他们所知的真空管电子学所能达到的速度。对于一台模仿图灵的通用机的多用途计算机来说，还需要一件物理装置，其功能相当于一个适当大小的存储器。在图灵抽象的通用机的纸带上，要想从一个方格移到较远的另一个方格，就需要进行多次一次一格的费力移动。就图灵在1936年的目的而言，这并没有什么不可以，那些理论机器并不需要做任何实际的事情。然而，快速的电子计算机却需要一个快速的存储器，这就要求被储存在存储器中任何位置的数据一步就能得到，也就是说，存储器应当是随机存取的。[47]

20世纪40年代后期，有两种装置可以用作计算机的存储器，它们是水银延迟线和阴极射线管。延迟线是由一个充满液体水银的管子构成的，数据以声波的形式存储在水银里，声波在管子的两端来回反射。阴极射线管常见于今天的电视机和计算机显示器，数据可以作为一种图样储存在真空管的表面。这两种装置都有严重的工程问题，但幸运的是，对于EDVAC项目来说，埃克特已经在战争期间对延迟线进行了改进，使它可以用于雷达。然而，到了20世纪50年代初，阴极射线管却成为了首选的存储介质。

在谈到这一时期时，人们通常都把正在研制的新的计算机说成是体现了存储程序概念，这是因为需要执行的程序在历史上第一次被存储在计算机内部。但不幸的是，这一术语模糊了这样一个事实，那就是这些机器真正革命性的地方是它们的通用特征，而存储程序只是达到目的的一种手段。图灵和冯·诺依曼的观点在概念上是如此简单，而且在很大

程度上已经成了我们的理智倾向，以至于我们很难理解它当时到底有多么新。认识到像水银延迟线这样的一项新发明的重要性要远比认识到一种新的抽象思想的重要性容易得多。埃克特后来声称，他在冯·诺依曼之前就已经想到了所谓的存储程序概念。他的证据是一本备忘录，其中谈到了在合金盘或蚀刻盘上实现的程序设计自动化。这里没有任何地方提到通用计算机的概念，也没有谈及指令和数据共存于其中的灵活的大容量存储器，哪怕连一点暗示都没有。而把既已做出的伟大进展说成是存储程序概念则会引发这种混乱。[15]

当直到埃克特和莫齐利试图在他们的工作的基础上制造出一种商品时，他们与冯·诺依曼和戈德斯坦之间的争吵达到了顶峰。他们寻求ENIAC和EDVAC的专利权。他们对于EDVAC专利权的申请毫无进展，这是由于冯·诺依曼草案报告的散发使之处于不受专利权限制的状态。正像我们已经解释的，他们的确收到了一份ENIAC的专利证，但后来却被法庭宣布为无效。埃克特和莫齐利显然预见到了通用电子计算机的商机，但却无法从这种先见之明中获利。[16]

随着埃克特和莫齐利的退出，摩尔学院失去了大部分优势，冯·诺依曼和戈德斯坦开始在普林斯顿高等研究院研制一种阴极射线管存储器计算机。冯·诺依曼开始寄希望于美国无线电公司所研制的一种专用真空管，但它并不管用，而英国工程师弗雷德里克·威廉姆斯（1911—1977）却设计出了能够使普通的阴极射线管有效地用作计算机存储器的一些方法，在许多年里，“威廉姆斯存储器”都一直独占鳌头。几台与研究院的机器类似的机器被制造出来了，它们根据约翰·冯·诺依曼的名字（Johnny）而被亲昵地称为“约尼阿克”（johnniac）。当IBM决定把通用电子计算机投入市场时，其第一种型号（701）与“约尼阿克”非常相似。[48]

阿兰·图灵的ACE（自动计算机）

当第二次世界大战结束的时候，英国国家物理实验所（NPL）进行了一次大规模的扩容，其中包括一个新的数学部门。J.R.沃玛斯莱（1907—1958）被任命为该部门的负责人，他已经意识到图灵1936年的早期论文“论可计算数”的实用含义。1938年，他甚至承担了用电子继电器设计一台通用计算机的任务，但后来放弃了这个想法，因为他发现这

样一台机器的速度将会非常慢。在1945年2月去美国访问期间，他看到了ENIAC，并且得到了一份冯·诺依曼的EDVAC报告的副本。他对此的反应是决定雇请阿兰·图灵。

到了1945年末，图灵已经完成了他那篇著名的ACE（自动计算机[Automatic Computing Engine]）报告。ACE报告与冯·诺依曼的EDVAC报告之间的比较表明，后者“是一份尚未完成的草案……更重要的是……它是不完整的……”但ACE报告却“是对计算机的一次完整描述，一直到逻辑电路图”，甚至还包括“预算为11200英镑”。在ACE有可能处理的10个问题的清单中，图灵凭借他那宽广的视野，把两个并非与数字资料直接相关的问题——下棋和简单的拼图玩具——包括了进来。[17]

图灵的ACE是一种与冯·诺依曼的EDVAC非常不同的机器，它们与这两位数学家的不同态度密切相关。尽管冯·诺依曼希望他的机器成为真正“通用的”，但他却把重点放在了数值演算上，且EDVAC（以及后来的“约尼阿克”）的逻辑构造都是为了在这个方向上有所推进。由于图灵发现ACE不适于做繁重的算术计算，所以ACE是沿着小得多的方式进行设计的，它更接近于“论可计算数”这篇论文中的图灵机，算术运算将通过程序设计（也就是通过软件而非硬件）来进行。为此，ACE的设计提供了一种特殊的机制，能够把预先编好的程序操作包含在一个更长的程序中。[18] 有人曾经建议沿着冯·诺依曼的方向改进ACE，图灵对此表现得极为刻薄：

> 它与我们的研制路线极为抵触，它更多地沿袭了美国的传统，即通过更多的设备而非思想来解决问题。……不仅如此，某些我们认为比加法和乘法更为基本的操作被漏掉了。[19]

图灵的极小主义思想注定对计算机的发展产生不了什么影响。但事后反思一下，我们就可以发现，ACE设计预示了所谓的微程序设计，它使程序员可以直接进行最基本的计算机操作。而且，我们今天所使用的个人电脑是建立在硅微处理器的基础之上的，它实际上就是芯片上的通用计算机，它们已经变得越来越精致了。而被一批计算机制造者所采纳的相反的纲领，即所谓的RISC（精简指令集计算机）结构，则在芯片上使用了最小指令集，它又一次与ACE的哲学相一致。

1947年2月20日，图灵在伦敦数学会上做了关于ACE和一般数字电

子计算机的讲演。他在开篇谈到了他在1936年所写的“论可计算数”的论文：

> 我考虑了这样一种机器，它有一个中心装置以及包含在一条无限长的纸带上的无限大的存储器……我的一个结论是，“经验”过程和“机械过程”这两种说法讲的是同一个意思。……像ACE这样的机器也许可以被看成我正在思考的那种机器的……实际版本。……至少非常相近。……像ACE这样的数字计算机……实际上是通用计算机的实际版本。[20]

图灵接下来提出了一个问题，即“一台计算机原则上可以在多大程度上模仿人的活动”。这使他提出，一台经过编程的计算机也许可以进行学习和犯错误。“有几条定理说的几乎就是……如果指望一台机器不会犯错误，那么它就不可能有智能……但对于一台不以不可错自居的机器能够表现出多少智能，这些定理却什么都没有说。”这间接提到了哥德尔的不完备性定理，我们将在下一章中更详细地加以讨论。图灵在讲演的最后请求“要让计算机进行公平竞赛”，这些计算机与人一样会犯错误，并且建议可以从下棋开始检验。要知道，所有这些想法都是在没有一种设备完成的情况下提出的！报告做完，听众们惊讶得鸦雀无声。[21]

当布莱奇利庄园的领导者们正在为获得充足的资源和支持发愁的时候，他们给温斯顿·丘吉尔写了一封信，丘吉尔立即保证他们能够得到所需的东西。ACE的建造不可能优先，而且NPL的行政官员也表现得相当糟糕。出色地研制了“巨人”的T.弗劳尔斯本该是研制ACE的理想人选，但他虽然与NPL签约，也做了一些计算机存储器延迟线的研制工作，却因忙于战后的电讯工作而只好作罢。这里面有ACE的极小主义设计的考虑，也许还夹杂着这样一种心情，那就是在技术方面可以信任美国人，而不要信任一个古怪的英国教师。在许多年里，这位教师在帮助赢得战争方面所做的事情一直都属于高度机密。当威廉姆斯表明他的阴极射线管存储器可以工作时（第207页），他被邀请加入研制ACE的工作，但他拒绝了。NPL天真地想象威廉姆斯会被雇来研制NPL计算机，但威廉姆斯已经获得了充足的资源，他可以在曼彻斯特研制他自己的计算机。最终，图灵离开了，他首先得到了剑桥的研究人员基金，然后接受了曼彻斯特大学提供的一个职位，他的老朋友兼战争期间的同事马科斯·纽曼正在在那里主持一个新的计算机项目。后来，一台小型的ACE

在NPL研制成功，它被称为“气行者ACE（Pilot ACE）”，并且一直良好运行了许多年。

埃克特、冯·诺依曼和图灵

这个通常被称为存储程序概念的故事通常有三个版本。第一种版本把这个概念看作冯·诺依曼的EDVAC报告所体现的天才的产物。埃克特对此大喊“不公”，他坚持说，早在冯·诺依曼加入摩尔学院的小组之前，他就曾建议研制存储程序计算机。他宣称，EDVAC报告是小组共同的想法。各种出版物都支持埃克特的立场。[22] 图灵的名字并没有被提到。戈德斯坦支持冯·诺依曼的说法，而没有考虑图灵的作用，他写道：

> 据我所知，冯·诺依曼是第一个清楚地懂得计算机本质上执行的是逻辑功能，而与电有关的方面则是辅助性的。[23]

当然，图灵对此也非常清楚。

ENIAC和通用计算机所体现的思想之间的鸿沟是如此之大，以至于很难相信埃克特会理解后者。当图灵抱怨“美国的传统是通过更多的设备而非思想来解决问题”的时候，他心里想到的很可能就是ENIAC。从图灵的“‘经验’过程和‘机械过程’这两种说法讲的是同一个意思”的结论可以很清楚地看出，把数从十进制转变成二进制，然后再转变回来，这是机器操作中最微不足道的事情。埃克特和莫齐利并没有认识到这一点，他们关心的是用十进制表示的输入输出量，他们研制了一台所有内部操作都用十进制来表示的庞然大物来解决他们的问题。人们遇到的许多具体问题都需要为微积分的某些极限运算找到近似值。由于被称为“差分分析器”的模拟机含有可以计算这种近似值的特殊器件，所以埃克特和莫齐利便把执行类似功能的组件用在了他们的ENIAC中。但是对于一台数字机器来说，这既不必要也不适当。微积分课本上讲述了计算这些值的方法，它需要的是对不超出算术的四种基本运算。

埃克特的确完成了一项与EDVAC有关的重大任务：他认为水银延迟线可以解决大容量存储器的问题。埃克特曾经研究过这些用于雷达的延迟线，所以对它们很是了解。因此，在那本他后来用于证明是他最先

想到了存储程序概念的备忘录中，据说他谈到了在合金盘上实现的程序设计自动化，而没有提到他了解很多的、更适于做存储介质的延迟线。

冯·诺依曼把计算机程序设计看作一种活动，把这种看法与图灵的看法做一番比较是很有趣的。冯·诺依曼称之为“编码”，并且明确指出它是一种办事员的工作，基本不需要理智参与。有一件逸事颇能发人深省，高等研究院让学生用人工把人类易读的计算机指令翻译成机器语言，有一个极富才华的年轻人提议写一个汇编程序来自动完成这种转换。据说冯·诺依曼曾经气愤地说，让一个有用的科学工具去做一项办事员的工作简直是浪费时间。在ACE报告中，图灵说计算机程序设计过程“应当是非常吸引人的，它不应被沦落为苦差事，因为任何非常机械的过程都可以交由机器本身去处理。”[24]

尽管故事的埃克特版本和冯·诺依曼版本今天仍然可以听到，但第三个版本已经变得越来越为人所知。这个版本说的是，冯·诺依曼从图灵的工作中得到了实际的通用计算机的思想。1987年，当我写文章阐述这种观点时，我发现持此看法的仅我一家。[25] 从那时起，人们越来越多地知道了图灵在战争期间对破译德国通信方面所起的作用，许多人也知道了他曾因同性恋风波而受到无耻的迫害。在伦敦和百老汇上演的《解密》（Breaking the Code）一剧戏剧性地描述了这些事情以及图灵的数学思想的重要性，它也是美国公共广播公司播出的一部电视剧的蓝本。[26] 电视纪录片也讲述了他的故事。于是，你瞧！阿兰·图灵的名字被列入了《时代》杂志（1999年3月29日）评选的20世纪最伟大的20位科学家和思想家名录。文章这样说：

> 现代计算机的创造是由如此众多的思想和技术进步共同完成的，以至于把它的发明完全归功于一个人是太不明智了。但有一个事实仍然成立，那就是每一个敲击键盘、打开电子数据表或Word处理程序的人都是在现实的图灵机上工作的。

的确如此！下面是《时代》杂志对冯·诺依曼的评论：

> 事实上，从耗资1000万美元的超级计算机到今天的无线电话和菲比（furbies）玩具上所使用的微小芯片，所有计算机都有一个共同点：它们都是“冯·诺依曼机”，都是冯·诺依曼基于图灵在20世纪40年代的工作所提出的基本计算机结构的变种。

知恩的国家对其英雄的回报

1948年秋，当图灵抵达曼彻斯特时，战争所带来的创伤还在愈合，一些住宅区还保留着工业革命早期的城市所具有的可怕外表。一位作家曾经引用了弗里德里希·恩格斯的一本名著，它评价了1844年曼彻斯特工人阶级住宅的惨状：

> 他[恩格斯]……所描述的情况……是在极为悲惨、恶化、残酷和非人道的普遍背景下发生的，类似的情况在地球上从未出现过。……一到这些院子，他便发现自己被“污垢和令人作呕的垃圾所包围，类似的情况再也不会找到。……毫无疑问，这是迄今为止我所见过的最可怕的住宿条件……在这里的一个大杂院中，正好在入口的地方……就是一个没有门的厕所，非常脏，住户们出入都只有跨过一片满是大小便的臭气熏天的死水洼才行。”[27]

当然，在接下来的一个世纪里，大众的卫生设施已经得到了明显改善，而且无论如何，在图灵的社会阶层中也有人可以不住在工人阶级的住宅区了。然而，图灵与一个“下层”成员之间的结交还是导致了灾难。

我们可以想想NPL的糟糕管理会使图灵感到多么沮丧，这不仅浪费了他的才智，还使他在ACE报告和向伦敦数学会所做的讲演中显露出来的雄心勃勃的梦想毁于一旦。与此同时，计算机正处于研制过程中。在剑桥大学，莫里斯·威尔克斯（1913—2010）领导了EDVAC类型计算机的研制，它被称为EDSAC。与图灵在NPL的处境不同，威尔克斯拥有充足的经费来完成这个项目。每当图灵想起自己曾在NPL嘲笑过威尔克斯的一本备忘录，说它属于那种“通过更多的设备而非思想来解决问题的美国传统”时，他定是懊恼万分。到了1949年，EDSAC开始运行，并且可以买卖。威尔克斯和他的合作者的所谓发现，即微程序设计以及对子程序的系统应用，在图灵的ACE报告中都被清楚阐明这只能加剧他的沮丧。在曼彻斯特，图灵本来据说是要在那里领导计算机项目的，但威廉姆斯明确指出，他对某位数学家关于他的计算机的研制方面的想法不感兴趣。也是在1949年，成功运行的Mark I型曼彻斯特计算机出色地证明了威廉姆斯把现成的阴极射线管用作存储设备的技术，这一技术不久就用在了美国的计算机上。然而，其基本逻辑设计乃是源自冯·诺依曼的EDVAC报告，而不是源自阿兰·图灵。[28]

关于图灵的ACE，赫尔曼·戈德斯坦曾经评论说，尽管这个设计“在某些方面有些吸引力”，但它“最终并没有站稳脚跟，而是被自然选择所淘汰。”[29] 说这是某种自然选择的结果实在是不公平的。体现了图灵思想的“飞行者ACE”运行得相当好。如果各种组织和资源都准备得很好，没有理由认为ACE计算机就工作不好。这里最好是在更为一般的背景下来理解，即计算机的哪些功能应由硬件来实现。哪些应由软件来实现。图灵曾经提出过一种相对简单的机器，它的许多功能都是由软件来实现的。但反过来，程序员要对基本的机器操作进行许多控制。这特别有利于编写旨在完成逻辑演算而非数值演算的程序。随着这个领域的发展，人们继续围绕对不能兼顾因素的这一权衡争论不休，最近的一场争论涉及RISC结构。[49]

当图灵1948年来到曼彻斯特大学时，尽管他仍然担任政府的顾问，但几乎没有人知道他在战争期间所做的事情。他之所以被雇用，是因为大家本以为他会在与威廉姆斯的Mark I计算机有关的事情上起一些领导作用。但后来事情的发展表明，工程师只顾忙于自己的事情，图灵在这些方面的想法并没有被执行得很好。他实际上并没有利用他的职位去介绍ACE报告中的那些优美思想，以使程序员的工作轻松愉快，而是成了直接与机器语言0和1打交道的使用者。他研究了一些他在战前曾经思考过的计算问题，但他的兴趣很快就转到生物学上去了。他试图回答生物是怎样从完全相同的细胞一步步地发展成为自然界中形态各异的生物的。由这个形态发生问题可以导出微分方程，图灵很自然地转向计算机来寻找关于这些方程的解的信息。虽然他用机器寻求的正是他曾经要求超越的那种数字压缩，但他在一些通俗文章和公众讲演中却表达了他长期以来对计算机拥有类似于人的智能的潜力的设想。

1951年圣诞节前夕，图灵设法与一个19岁的年轻人阿诺德·默里开始了短暂的暧昧关系。默里来自一个穷苦的工人阶级家庭，是一个非常聪明的年轻人。当图灵与他在街上搭话时，默里正处于缓刑期，他曾经在一次小偷小摸中被捉。图灵邀请他到自己的住处，在默里眼里那儿一定是豪华住宅。圣诞节过后不到一个月，图灵有一天晚上回到家时发现他的住所失窃了。尽管丢失的东西总共价值还不到50英镑，但图灵还是非常心烦意乱。是默里想到了是谁实施了行窃——是他认识的一个名叫哈瑞的人。哈瑞显然认为对一个同性恋者实施盗窃是绝对安全的，因为他不敢去报警。一个谨慎的人面对阿兰的处境时当然不致愚蠢到去报警，但图灵却偏偏去了警察局。

警察局没费多少气力就弄清了图灵和默里之间的事情，当被讯问时，图灵一切都没有隐瞒。他不认为自己在性方面的情感有何可耻之处，也不认为他实现这些情感的无害手段有何过错。但尽管如此，法律的规定是很清楚的：图灵和默里彼此愉悦对方的所作所为是“极为下流的”，最多可以被判处2年徒刑。审理图灵案件的法官出于他所谓的仁慈动机，决定如果图灵同意进行1年的激素注射治疗以减少性冲动，那么他就可以免于坐牢。所使用的激素是雌性激素，无论它会给图灵的性冲动带来什么影响，它都会附带地使他的胸部发育。

1938年10月，阿兰看了沃尔特·迪士尼的《白雪公主和七个小矮人》。“他被一个场面深深地吸引住了，邪恶的女巫正把一个系在线上的苹果摇摇晃晃地放入一锅滚开的毒药中，口里还念念有词：

> 毒液浸透苹果，
>
> 沉睡般的死亡也随之穿透。”

他似乎喜欢一遍又一遍地念叨着这几句话。[30] 1954年6月7日，阿兰·图灵咬了半个在氰化物溶液中浸泡过的苹果，结束了自己的生命。是什么导致他做出了这个不可逆转的行为呢？人们对此有种种猜测。《解密》一剧说，在他被定罪之后，政府当局曾经反对他出国旅行，因为这成了他的性伴侣的最好来源。在英国，性也许已经变得太过危险以至于不能尝试。在20世纪50年代的气氛中，当局的确反对他出国旅行，这似乎是不争的事实。被定罪之后，官方不再允许他阅读秘密文件，但无法抹去的是他大脑中所携带的秘密。不过可以肯定的是，他在一次去挪威的旅行中遇上了一个男人，这个人曾经被警察局阻止，并且当他到英国探访图灵时被驱逐出境。唉，阿兰·图灵很可能就这样被一个他曾经为之做出过如此巨大的贡献（虽然未获赏识）的国家的政府当局逼上了绝路。

第9章　超越莱布尼茨之梦

图灵在伦敦数学会的演讲中说：

> 我希望数字计算机能够最终激起人们对符号逻辑的极大兴趣……人与这些机器进行交流的语言……构成了一种符号逻辑。[1]

图灵所暗示的逻辑和计算之间的关联是本书的一个基本主题。然而，读者们也许仍然会问：逻辑和计算之间是如何彼此关联的？算术与推理有什么样的关系？动词“推断”（reckon）的通俗用法为我们提供了一条线索，在这里它没有了通常的含义：“演算。”

> 我推断他此时正在月光下对她说着甜言蜜语。

我们正在听一部B级电影中忧郁的男主角谈论他的情敌，他（和我们一样）并不知道赢得她的芳心的正是我们的男主角。他在言谈中未曾想到算术，他只是在谈论推理。他基于自己自认为知道的情敌的那些背信弃义的做法进行推理。“推断”一词的用法暗示了推理和演算之间真实而深刻的关联。用数进行推断本身就是一种推理形式，人们所做的大量推理都可以被看成一种计算。就像我们的例子所显示的，这种关系通常都能被人理解（至少是在一种下意识的层次上），这真是很有趣。又如，当某个人被说成是一种善于算计的人的时候，我们对这种评价完全心领神会。

把逻辑推理归结为形式规则的努力可以追溯到亚里士多德。他的工作是莱布尼茨关于一种普遍的计算语言的梦想的根本基础。图灵说明了所有的计算都能够在他的通用机上完成，这一成就也是以亚里士多德的工作作为支撑的。计算和逻辑推理确实是同一个硬币的两面，利用这一洞见，我们不仅可以为计算机编程，从而使其能够完成各种任务，而且还能设计研制计算机。

1999年1月，美国科学促进会在加利福尼亚的阿纳海姆召开了年会。这次会议的主题为“新世纪的挑战”。这是一次严肃认真的内容丰富

的会议，它就大量课题举行了演讲和讨论会。当我浏览议程表时，一个名为“逻辑在计算机科学中的非凡效力”的专题讨论会引起了我的注意。五位领衔的计算机科学家（也是逻辑学家）在发言中证实了图灵预言的正确。在该讨论会的会议议程上，他们写道：

> 在过去的25年里，逻辑和计算机科学之间发生了广泛而持续的相互作用，而且仍在不断加强。在许多方面，逻辑为计算机科学提供了一种统一的基础框架和建立模型的工具……这次讨论会通过展示逻辑在其中扮演着至关重要的角色的一些计算机科学领域，如人工智能、计算复杂性、分布式计算、数据库系统和设计验证等等，从而为逻辑在计算机科学中的惊人效力进行一次全面考察。[2]

计算机、大脑和心灵

1965年，麻省理工学院的计算机科学家约瑟夫·外岑鲍姆编写了一个名为ELIZA的程序，用户只要通过键盘输入英语句子，就可以与之对话。ELIZA有意以一种相当粗糙的方式去模仿一个启发式的精神治疗师。以下是我和ELIZA之间一次真实对话的记录：[3]

> ELIZA：你好。请告诉我你的烦恼。
>
> 马丁：我的烦恼很难描述。
>
> ELIZA：为什么你的烦恼难以描述？
>
> 马丁：因为每当我试图去描述它时，我就会大叫。
>
> ELIZA：这是真正的原因吗？
>
> 马丁：也许不是。
>
> ELIZA：你似乎并不很确定。
>
> 马丁：啊，确定。
>
> ELIZA：你能仔细地讲讲吗？

马丁：可以。但这将涉及一些个人隐私。

ELIZA：你为什么如此确信？

马丁：我试图维护我的尊严。

ELIZA：你的尊严？

马丁：你不会理解这些的。

ELIZA：我们是在讨论你，而不是我。

马丁：我已经受够了。再见！

ELIZA：你说你已经受够了。

马丁：我说再见。

ELIZA：你能说得再详细一些吗？

马丁：再见。

ELIZA：再见。请付75美元。

当然，ELIZA什么都不理解。该程序执行的是一种非常简单的算法，如果句子中出现关键词，那么它就会做出响应，否则就只是重复用户的句子。例如，“你”是一个关键词，ELIZA对它的响应通常是“我们是在讨论你，而不是我。”句子“你是我的母亲”中包含了“母亲”和“你”这两个关键词。由于“母亲”这个关键词的级别要比“你”这个关键词高，所以ELIZA对于该句的响应将是“再告诉我一些你家庭的情况。”“可以”这个关键词引起了“你为什么是如此肯定？”的响应。注意，只有当“再见”一词出现在句首时，ELIZA才会对它做出响应。

1950年，阿兰·图灵发表了他的经典论文“计算机与智能”。他预言，世纪末将会出现这样一种灵巧的人机对话程序，人在对话时将不能分辨他是在与一个人还是一台计算机交谈。[4] 他的预言是错误的：如今，那些声称能对普通的英语句子做出响应的交互式程序远比ELIZA复杂，但它们之中最好的也远远赶不上一个普通的5岁小孩的语言能力。

图灵其实是在不陷入哲学和神学问题的泥淖的情况下，探讨一台计算机是否可以表现出智能行为。为此，他提出了一种客观的、易于操作的检验方法：假如一台经过编程的计算机能够与一个有着正常理性的人谈论任何话题，而且用户判断不出来他是在同一个人交谈还是在同一台机器交谈，那么我们就可以说这台计算机是具有智能的。然而，我们距离能够编出这样的计算机程序还相当遥远，而且许多人对于这种判定计算机是否具有智能的方法并不信服。

正当计算语言学家们为计算机运用日常语言的能力而不懈奋斗之时，人们也自然会在那些不依赖于日常语言的领域中去探求机器智能。其中有一个领域就是下棋。当一个人国际象棋下得不错时，很难否认他表现出了智力思考。众所周知，我们已经有了能够把国际象棋下得很好的程序，许多普通的玩家必须选择该程序中的较低级别，这样才不会被经常击败。1996年2月，电脑“深蓝”成功地击败了国际象棋世界冠军加里·卡斯帕罗夫。那么，我们能说“深蓝”具有智能吗？哲学家约翰·R.塞尔在一篇文章中以他那惯用的挑衅风格告诉我们，严格说来，“深蓝”甚至不能被认为是在下棋：

> 以下我们将说明“深蓝”内部发生了什么。这台计算机有一串被程序员用来表示棋盘上棋子位置的毫无意义的符号，还有一串表示可能的移动的同样毫无意义的符号。这台计算机并不知道这些符号就表示棋子和它们的移动，因为它什么都不知道。[5]

为了把这一点讲清楚，塞尔把他那个著名的比喻稍稍改造了一下。原来的故事说，屋子里有一个人从屋外收到了一些符号，然后他通过查一本书来决定他将发回哪些符号。那本书是以一种特别的方式写成的，这使得来回往复的符号构成了一次中文对话。但这个人并不懂中文，他不明白那些符号代表着什么。我们且不考虑从这个奇异的故事中可以得出什么结论，让我们转到塞尔的“象棋屋”：

> 想象一个不晓得如何下国际象棋的人被锁在一间屋子里，在那里他被给予了一组在他看来毫无意义的符号。他不知道这些符号代表着棋盘上的位置。他在一本书中查寻他该怎样做，然后送回一些无意义的符号。我们可以假定，如果这本书（也就是那个程序）编写得很是巧妙，那么他将赢得象棋比赛的胜利。屋外的人则会说：“这个人懂得国际象棋，他的确是一个象棋高手，因为他赢

> 了。”但他们完全错了。这个人根本不懂国际象棋，他只是一台计算机。这个比喻的要害在于：假如这个人不懂国际象棋，而只是按照下棋程序行事，那么任何其他按照下棋程序行事的计算机也不懂象棋。

读者们也许已经注意到了，在这个例子中硬件和软件是分离的。屋中那个人的功能其实就相当于一台粗制的通用计算机。当然，一台未安装软件的计算机是不可能玩国际象棋的，只有当那个人有一本指令书时，下棋才能进行。下面是塞尔的比喻的我自己的版本：

> 有一个早熟的孩子，他的母亲对国际象棋充满热情。他厌倦了只是在一旁观看他的母亲下棋，于是就要求他的母亲允许他和母亲下盘棋。他母亲同意了，但前提是只有当她说可以移动棋子时才移动棋子，而且要将棋子放在她所指定的位置。他按照要求做了。母亲在他耳边悄悄说什么，他就做什么。最终，他把对方将死了。看到这一幕情景，塞尔说那个孩子一点也不懂国际象棋，而且他肯定也不是在下棋，谁又能不同意呢？

当代哲学家所使用的方法之一就是讲一些荒谬的故事，从而使那些并不明显的联系显示出来。但是建造现实的“象棋屋”也并非毫无意义。我有一个同事，他曾是“深思”设计小组的成员。“深思”是一台功能强大的下棋计算机，它是“深蓝”的前身。他提供给我一些数据，在此基础上我演算出，假如构成“深思”的硬件和软件被表示成一本由人可以执行的指令写成的书（更像是一个图书馆），那么把棋子移动一步就要花去好几年的时间。于是，最好是把一家人放在那个象棋屋里，当双亲去世后，孩子们还可以接着干。否则的话，一局比赛也不可能完成。

塞尔告诉我们说，“深蓝”“有一串毫无意义的符号”。然而，假如你能在“深蓝”运行时看到它的内部，那么你将看不到任何符号，无论是有意义的还是无意义的。电子正在电路中来回运动。这就好像，假如你能在卡斯帕罗夫下棋时看到他的头骨内部的情况，那么你将看不到任何棋子，而只能看到神经元的脉冲。我们的大脑如何组织才能处理那些所谓的符号信息，人们对这个问题直到现在也是知之甚少。而电脑（比如“深蓝”）为这个目的是如何被组织起来的，人们就清楚多了，因为毕竟是工程师和程序师制造了它。但是在这两种情况下，在大约分子层次上起作用的过程都被整合在样式中了，而认为样式与符号处理有关是有

帮助的。塞尔告诉我们说，“深蓝”中的符号是毫无意义的。所以卒和马到底“意谓”着什么，这并不是一个有用的问题。

塞尔强调了“深蓝”不“知道”它在下棋这个事实。事实上，他坚持说“深蓝”不知道任何东西。而那些富有专业知识的工程师却有可能声称，“深蓝”的确知道各种东西。例如，它知道能将给定方格中的棋子移动到哪几个方格去。这完全取决于“知道”是什么意思。尽管如此，我们可以说“深蓝”并不知道它在下棋。但我们能因此就下结论说它事实上不在下棋吗？这里是另外一个比喻：

> 人类学家对新几内亚北部的克斯鲁普人的研究导致了一个非同寻常的发现，它肯定是人类有史以来最为惊人的巧合之一。虽然克斯鲁普人在被发现以前一直过着与世隔绝的生活，但他们却举行这样一种宗教仪式，他们两两组合参加一种完全等同于我们的国际象棋游戏的符号性的仪式。他们并没有使用棋盘或棋子，而是在沙地上设计了精巧的格子。首先发现克斯鲁普人的人类学远征队的领导者斯普兰蒂德博士本人是一个热情的业余棋手，正因为此，他才在沙地上所画的图案中看出了与国际象棋比赛中的移动完全等价的东西。

这些克斯鲁普人是在下国际象棋吗？他们肯定不知道自己下的正是国际象棋。塞尔也许会回答说：“但克斯鲁普人是有意识的，而‘深蓝’没有。”一个由程序控制的计算机是否会有意识，这在塞尔和其他人关于这些问题的讨论中扮演着主要的角色。无论未来的情况如何，我们当然必须得同意“深蓝”没有意识。

我们的意识是我们体验自身独特的个体性的一种主要方式，但我们只是从内部才知道它的。我们可以经验到自己的意识，却不能经验到他人的。我把我的意识经验为一种内在的交谈。我的妻子向我保证，她的意识是由视觉图像所支配的。她的意识和我的意识真是同一种东西吗？它到底是什么，为了什么目的而服务？当我写作时，我试图寻找合适的词，（如果幸运）它从我意识的深处浮现了出来。我不知道我的大脑是如何做到这一点的。时至今日，意识的现象依旧是那么神秘。

出于很好的理由，图灵和冯·诺依曼都把电脑与人脑进行了比较。既然人们有那么多不同的思维模式，他们猜想，我们之所以能够做那么

多极为不同的事情，是因为在我们的大脑中嵌入了一台通用计算机。这就是为什么当冯·诺依曼着手设计EDVAC时，他会被人工神经元理论深深震撼。通用计算机所能做的只是执行算法。塞尔说：“人类很少做那种单纯的计算，我们很少把时间花在执行算法上。”他真的如此确信吗？当被问到这样一个问题：你读过查尔斯·狄更斯写的东西吗？答案（有或没有）一下子就从脑海深处冒了出来。对此我们是如何做到的？我们不知道。可以假设这是通过某种算法过程从我们大脑中的数据库里提取所需的信息而实现的，这种假说至少从表面看是相当吸引人的。对于计算机如何处理从电视摄像机输入的原始的视觉资料的研究，将在很大程度上帮助我们理解，我们的大脑是通过什么样的过程把视网膜传入大脑的原始资料转变为呈现给我们的清晰图像的。我们不知道这些事情是否就是通过我们的大脑执行算法来实现的，但我们并不能肯定这就不是它的实现方法。

罗杰·彭罗斯是一位杰出的数学家和数学物理学家，他在宇宙的几何结构方面做出了激动人心的工作。他也曾考虑过人类心灵的运作是否从本质上说是算法的这一问题。彭罗斯援引哥德尔的不完备性定理，明确地回答说：不是。以下是一种表达哥德尔定理的方式：

> 给定一种能够逐条产生关于自然数的真陈述的算法，那么我们总是能够得到另一条关于自然数的真陈述（我们称之为哥德尔句子），它是不能通过该算法生成的。[6]

彭罗斯论证说，没有一种声称等价于心灵活动的特殊算法能够满足需要，因为心灵总有可能通过一种“洞察”活动而发现哥德尔句子对于该算法为真。这则论证是极端错误的，至于其原因，图灵早在1947年为伦敦数学会所做的演讲中就已经解释过了，那次演讲足足比彭罗斯所写的文章早了40年。图灵指出，哥德尔定理仅仅适用于那些只生成真陈述的算法。但没有一个数学家能够声称不会出错。我们都会犯错误！因此，哥德尔的定理无法防止把人类心灵的数学能力等同于一种既能产生真语句又能产生假陈述的算法过程。[7]

塞尔和彭罗斯拒不承认人类的心灵就其本质而言等同于一台计算机。但他们两人都心照不宣地接受了这样一个前提，即不论人类的心灵可能是什么，它都是由大脑产生出来的，都服从物理化学定律。而库尔特·哥德尔则愿意相信，大脑实际上就是一台计算机，但他拒不接受超

越于人脑的心灵并不存在的观点。事实上，古典的心—身问题是哥德尔所关注的问题的核心。他认为心灵以某种方式独立于我们作为物理实体的存在，他的这种立场通常被称为笛卡儿的二元论。[8]

这里的讨论已经远远超出了莱布尼茨之梦，它把我们带到了一个介于哲学和科幻小说之间的地方。的确，只要注意到自EDVCA和ACE报告出现以来的计算机的发展状况，我们在预测它们未来能够做什么和不能做什么时，便很有理由要谨慎一些。

尾声

我们已经领略了一批卓越的革新者横跨3个世纪的生活。他们中的每一位都以这样那样的方式关注人类推理的本性。他们每个人的贡献加在一起便构成了理智的母体，由此孕育出了通用数字计算机。除了图灵，他们之中没有一个人意识到自己的工作可以被如此应用。莱布尼茨看得很远，但还没有远到这种程度。布尔几乎不可能想到，他的逻辑代数会被用于设计复杂的电路。如果弗雷格发现与他的逻辑规则等价的东西会被纳入实现演绎的计算机程序，他定会大吃一惊。康托尔当然从未料到他的对角线方法会产生出来什么样的结果。希尔伯特用于确保数学基础的纲领被引向了一个非常不同的方向。即便是一直过着心灵生活的哥德尔，也几乎没有想到自己的工作可以在机械装置上得到应用。

本书中的故事强调了观念的力量以及预测它们结果的徒劳。汉诺威的公爵们认为他们知道莱布尼茨最应当做什么：编写他们的家族史。今天的情况更是如此，那些为科学家们提供生活和工作所必需的资源的人竭力要把他们引向那些被认为能够尽快出结果的方向。这不仅在短期内可能是徒劳的，而且更重要的是，如果只重眼前利益，而轻视那些不会带来直接回报的研究，那么最终遭殃的还是我们的未来。

注释

所有括号中的文献均见于后面的参考书目。

引言

[1] 引文出自[Ceruzzi]，第43页。霍华德·艾肯（1900—1973）组建了哈佛大学计算实验室，并于20世纪40—50年代在哈佛大学帮助设计和建造大型计算设备。

[2] 引文出自一篇向伦敦数学会所做的讲演[Turing 1]，第112页。阿兰·图灵是本书第7章和第8章的主题。

第1章

[1] 关于莱布尼茨的传记材料，我主要取自[Aiton]。

[2] 关于莱布尼茨的*Dissertatio de Arte Combinatoria*（唉，原文为拉丁文），参见[Leibniz 1]。

[3] 莱布尼茨在巴黎的数学著作在[Aiton]中有所讨论，更详尽的见于[Hofmann]。

[4] 引自[Leibniz 3]。

[5] 莱布尼茨关于推理和解方程机器的著作，参见[Couturat]，第115页。

[6] 对牛顿和莱布尼茨所创立的微积分的数学细节感兴趣的读者将会喜欢[Edwards]中的细致讨论。关于微积分历史发展的优秀论述，读者还可参考[Bourbaki]，第207—249页。

[7] 关于莱布尼茨的微积分运算还有另一个有趣的故事（但它其实应归入另一本书）：他对无穷小量的使用。无穷小量被认为是非常小的正数，以至于无论这个数加到自身多少次，数1（甚或0.0000001）也永

远达不到。这些量的合法性从一开始就受到了挑战；哲学家贝克莱大主教嘲笑无穷小量是“已经死去的量的幽灵”。到了19世纪末期，数学家们都认为对无穷小量的使用的合理性不可能得到证明（尽管物理学家和工程师继续使用它们）。关于莱布尼茨所使用的无穷小方法以及20世纪逻辑学家亚伯拉罕·鲁宾逊对它们的最终肯定，参见[Edwards]。《科学美国人》上的文章[Dav-Hersh 1]对鲁宾逊的成就给出了另一篇论述。

[8] [Aiton]，第53页。

[9] [Mates]，第27页。关于这些著名女性的更多内容以及莱布尼茨关于女性理智能力的信念，也参见此文献第26—27页以及其他文献。

[10] 引用的这封致洛比达的信写于1693年4月28日（[Couturat]，第83页）。古杜拉的引文也出自这篇文献的同一页。关于“阿里阿德涅之线”，参见[Bourbaki]，第16页。

[11] 这封莱布尼茨论述其普遍文字的写给让·伽鲁瓦的信[Leibniz 2]写于1678年12月。译文是我从法语翻译过来的。

[12] [Gerhardt]，第7卷，第200页。

[13] [Parkinson]，第105页。

[14] 关于莱布尼茨的逻辑演算（这里只举了一个小例子），参见[Lewis]，第297—305页。莱布尼茨没有使用“=”，而是使用了∞。[Swoyer]这篇有趣的文章从20世纪的视角对这个系统进行了一次彻底重建。

[15] 关于对莱布尼茨试图超越亚里士多德的分析的某些讨论，参见[Mates]，第178—183页。

[16] [Huber]，第267—269页。

[17] [Aiton]，第212页。

第2章

[1] 关于莱布尼茨与卡罗琳娜公主的友谊及其与塞缪尔·克拉克的通

信的内容取自[Aiton]，第232页、第341—346页以及[Britannica]中的“卡罗琳娜（1683—1737）”以及“塞缪尔·克拉克（1675—1729）”词条。

[2] 关于乔治·布尔的传记信息主要取自[MacHale]。

[3] [MacHale]，第17—19页。

[4] “强烈的欲望和激情”：[MacHale]，第19页。

[5] [MacHale]，第30—31页。

[6] [MacHale]，第24—25页。

[7] [MacHale]，第41页。

[8] 代数定律中最重要的有加法和乘法的交换律：

x+y=y+x xy=yx

以及分配律

x（y+z）=xy+xz

我们使用了通常的代数约定，比如写xy而不写x×y。

[9] 两个微分算子的乘法（意思是先作用第一个再作用第二个）并不总是遵循交换律。

[10] 布尔的金质奖章：[MacHale]，第59—62页，第64—66页。除运用微积分方法的著作之外，布尔还在1842年的《剑桥数学杂志》上以两部分发表了一篇论文，可以认为这篇文章创立了一个新的重要的代数分支，布尔后来再也没有研究过不变量。我们将在大卫·希尔伯特一章中再次讨论不变量。

[11] 与布尔对待证明和极限过程的模糊态度不同，当时的欧洲大陆正努力为它们发展一种恰当的严格基础。有兴趣的读者可以参见[Edwards]，特别是第11章。

[12] 不要把苏格兰哲学家威廉·汉密尔顿爵士（Sir William Hamilton）与其同时代人——爱尔兰数学家威廉·罗万·汉密尔顿（Sir William Rowan Hamilton）爵士混淆起来。

[13] [Boole 2]，第28—29页。

[14] [Daly]，[Kinealy].

[15] [MacHale]，第173页。

[16] [MacHale]，第92页。

[17] [MacHale]，第107页。

[18] [MacHale]，第240—243页。

[19] [MacHale]，第111页。

[20] [MacHale]，第252—276页。

[21] x和y的交集的现代记法是x∩y而非xy。空集也通常用丹麦语字母Φ而非0来表示。当然，他使用的记号对布尔来说是重要的，因为这样可以使它很容易与普通的代数相联系。

[22] 布尔把+这一操作严格限于没有共同元素的类。这里我们遵循当前的用法，而不去加强这一限制。于是，x+y就是属于x或y或两者的事物的类。今天，我们说x和y的并集，写作x∪y。布尔还把x-y限制在y表示的类是x表示的类的一部分的情况。但这一限制也是没有必要的。

[23] [Boole 2]，第49页。

[24] 正如布尔所强调的，对三段论有效性的论证中涉及代数的是把1个变量从具有3个变量的两个联立方程中消去。

尽管布尔完全认识到，“所有X都是Y”这一形式的命题在他的代数中可以用X（1-Y）=0来表示，但他宁愿使用X=vY，其中v是他所谓的不定符号（indefinite symbol）。数学家查尔斯·格雷夫斯清楚地表明了这一点（[MacHale]，第70页）。这真是一个糟糕的想法，它给布尔的

体系造成了非常不必要的复杂。

[25] 布尔把二级命题与他的类代数相联系的方法是把时间引进来。事实上，布尔会把时间瞬间的类与每一个命题联系起来。要想说命题X为真，布尔会记为X=1，意指表示命题为真的瞬间的类包含了整个时间段。类似地，X=0将表示X为假，因为没有时间的瞬间是X为真的。假设命题X&Y表示X和Y均为真，它为真的瞬间的集就是交集XY。最后，要想让如果X，那么Y这一命题为真，就需要无论什么时候X为真，Y也必须为真，也就是说没有时间是X为真而Y为假的。因为X（1-Y）=0（[Boole 2]，第162—164页。）

[26] [Boole 2]，第188—211页。

第3章

[1] 关于罗素的信、弗雷格的回复以及罗素后来的评论，参见[van Heijenoort]，第124—128页。

[2] 2.关于弗雷格声名狼藉的日记以及迈克尔·达米特的评论，参见[Frege 2]。

[3] 我非常感谢莱比锡大学的洛塔·克莱泽尔教授和蔼地回复了我有关弗雷格信息的请求。尽管克莱泽尔教授主要把时间花在处理德国统一所带来的问题上，但他还是挤出时间回复了我的请求。[Bynum]中的特劳尔·拜纳姆简明的传记也很有帮助。

[4] 我发现[Craig]是一份出色的关于德国史的资源。关于第一次世界大战的起源，也参见[Geiss]，[Kagan]。弗雷格寄给哲学家路德维希·维特根斯坦的一些明信片保存了下来，战争期间，后者曾在奥地利军队中当一名炮兵观察员。毫不奇怪，它们表明弗雷格曾是一个爱国的德国人[Frege 1]。

[5] [Frege 2].

[6] 同上。

[7] [Sluga]，[Frege 1]，第8—9页。

[8] 关于所引的评论，参见[van Heijenoort]，第1页。同一资料第1—82页还包括了附带注解的弗雷格《概念文字》的出色翻译。另一个译本收在[Bynum]，第101—166页。

[9] 我们使用的这些符号是今天普遍使用的符号，而不是弗雷格所使用的那些符号。当然，基本的洞见是认识到什么需要被符号化，而不是所使用的是什么特殊符号。弗雷格的符号并没有被广泛接受，这部分是因为它们给排字工人造成了麻烦，但更主要的是因为经由伯特兰·罗素改造的意大利逻辑学家朱塞佩·皮亚诺所使用的记号更为人所熟知。

[10] 弗雷格写道："我想创立的不仅仅是一种'推理演算'，而且也是一种莱布尼茨意义上的'符号语言'（*lingua charactera*）。"引自[van Heijenoort]，第2页。也参见[Kluge]。

[11] 这条规则被称为"肯定前件推理"（*modus ponens*）。该术语来自12世纪的经院逻辑学家。

[12] 我们现在通常称弗雷格的逻辑为一阶逻辑。这是要把它与量词∀和∃既作用于属性又作用于个体的逻辑系统相区别。下面是一个二阶逻辑的句子的例子：

$$(\forall F)(\forall G)[(\forall x)(F(x)\supset G(x))\supset(\exists x)(F(x)\supset(\exists x)G(x))]$$

事实上，弗雷格的确考虑了属性的量化，从而超越了一阶逻辑，所以我们把一阶逻辑称为"弗雷格的逻辑"是不确切的。

[13] 严格说来，与弗雷格自己的陈述相比，这种对"数"的解释更近似于伯特兰·罗素的说法。但这更易说明它为什么容易落入罗素悖论。

[14] 当本书正在写作之时，又有一些有趣的工作显示，弗雷格用逻辑发展算术的相当一部分计划都可以被挽救[Boolos]。

[15] [Frege 3].

[16] [Dummett]，[Baker-Hacker].

[17] 为了清晰起见，能够精确地说出计算机程序设计语言中出现的惯用语的含义是很重要的，也就是说，为这样一种语言提供语义学。关于这个问题已经做过很多研究，有一种解决方案被称为指称语义学，它完全是基于弗雷格的思想。参见[Dav-Sig-Wey]，第465—556页。

第4章

[1] [Rucker]，第3页。

[2] 引自[Dauben 1]，第124页。由莱布尼茨的法文原文译出[Cantor 1]，第179页。

[3] [Dauben 1]，第120页。

[4] [Frege 4]，第272页。这里的引文出自弗雷格对康托尔某些工作的评论。本章后面还有关于这个评论的更多内容。

[5] 关于康托尔的传记信息，我参考的是[Grattan-Guinness]，[Purkert-Ilgauds]以及[Meschkowski]。

[6] [Meschkowski]，第1页（我的翻译）。

[7] 数学家和物理学家对于三角级数的极大兴趣是被激发出来的，它来自于法国数学家傅立叶在19世纪早期的惊人发现，即三角级数的收敛极限似乎很不明显。三角级数的一个例子是

$$\cos x+\frac{\cos 2x}{4}+\frac{\cos 3x}{9}+\frac{\cos 4x}{16}+\frac{\cos 5x}{25}+\cdots .$$

值得注意的是，如果x处于0和2π之间，则该级数收敛于$\frac{1}{4}x^2-\frac{1}{2}\pi x+\frac{1}{6}\pi^2$。（“角”x以弧度为单位。）如果x被定为0，则我们就得到

$$\frac{\pi^2}{6}=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots$$

就像莱布尼茨级数收敛于π/4一样，这个结果把π与自然数联系了起来，这里的自然数是完全平方数：

1×1=1，2×2=4，3×3=9，4×4=16，5×5=25，…….

[8] [Euclid]，第232页。

[9] [Gerhardt]，第1卷，第338页。拉丁文的译者是Alexis Manaster Ramer。

[10] 正如我们在小学所学到的，不同的分数可以表示同一个数，比如

$$\frac{1}{2}=\frac{2}{4}=\frac{3}{6}\ \cdots$$

所以分数与自然数之间的一一对应是作为符号的分数的对应，而不是符号所代表的数的对应。而这是很容易确定的：只要从分数序列中除去所有那些各项不是最小的分数就可以了。

[11] 超限数的存在已经被法国数学家刘维尔于30年前以一种完全不同的方式证明了。刘维尔所能够证明的是，一个小数展开包含许多极长的0串的数只能是超限数。刘维尔的方法可以应用的一个例子是

$$0.10100001\underbrace{000000000000000000000000000}_{27}1\underbrace{000\cdots010}_{64}\cdots$$

这里两个1之间的连续的0的数目是1^1 =1，2^2 =4，3^3 =27，4^4 =64等。当康托尔写这篇论文的时候，证明π是超越数尚待10年。而 $2^{\sqrt{2}}$ =2.665144142690225190是一个超限数的事实直到1934年才终获证明。

[12] [Grattan-Guinness]，第358页。

[13] 康托尔对于基数的记法在今天使用得不多了。现在我们写|M|而不是写 $\bar{\bar{M}}$ 。

[14] 事实上，对于任何两个不等的基数，其中一个必定要大于另一个这一命题对于无穷集来说并不是显然成立的。在康托尔活着的时候，这个问题并未真正得到澄清。

[15] 要想弄明白为什么由一切自然数集所组成的集合的基数与实数集的基数是相同的，考虑一下只包含0和1两个数的数值的二进制表示是有益的。当我们写1/3=0.3333……时，它的意思其实就是

$$\frac{1}{3}=\frac{3}{10}+\frac{3}{100}+\frac{3}{1000}+\frac{3}{10000}+\cdots$$

在二进制中，小于1的正实数可以用0和1的无限长的串来表示。例如

$$\frac{1}{4}=0.0100000000\cdots,$$

$$\frac{1}{3}=0.0101010101\cdots,$$

$$\frac{1}{\pi}=0.101000101\cdots,$$

$$\sqrt{\frac{1}{2}}=0.1011010100\cdots.$$

这里，当我们写下1/3=0.0101010101……时，它的意思其实就是

$$\frac{1}{3}=\frac{1}{4}+\frac{1}{16}+\frac{1}{16}+\frac{1}{64}\cdots$$

（分母是2的连续幂，而非10的连续幂。）

现在，无论以什么自然数集开始，我们都可以像下面这样找到唯一一个实数与之对应：如果n是给定集合的一个成员，那么我们就在第n个位置上写1，否则就写0，这样我们便产生了一串0和1的序列。例如，如果我们始自偶数集，那么我们得到的序列就是0.01010101……，即我们已经看到的1/3。如果始自奇数集，那么我们得到的序列就是0.101010……=2/3。

这就表明，由一切自然数集所组成的集合的基数与0与1之间的实数集的基数是相同的。但康托尔证明了（这其实并不困难）这个集合的基数与由所有实数所组成的集合的基数是相同的。

还有一点技术上的小麻烦我必须提及。某些有理数将有两种不同的二进制表示，所以它们将对应于两个不同的自然数集。例如：

$$\frac{1}{2}=0.1000000\cdots$$

$$=0.0111111\cdots$$

所以实数1/2既对应着仅由1组成的集合，又对应着由除1以外的所有自然数组成的集合。尽管这破坏了我们的一一对应关系，但利用有理数集的基数为$\aleph_0$的事实，这一困难便可克服。

[16] 正如康托尔所指出的，由一切实数集所组成的集合的基数也就是由一切从实数到实数的函数所组成的集合的基数。

[17] 参见[Grattan-Guinness]以及[Dauben 1]。Barbara Rosen博士曾就这个问题向我友好地提出了专业的建议。

[18] 感谢Michael Friedman曾就康德及有关事项提供的帮助（尽管他不应为我对黑格尔的攻击负责）。

[19] [Cantor 1]，第382页。

[20] [Frege 4]。非常感谢Egon Börger, William Craig, Michael Richter和Wilfried Sieg对这段话以及前一注释中的话的翻译。

第5章

[1] 关于希尔伯特的信息，我参考了[Reid]的传记、奥托·布鲁门塔尔所写的传记文章[Hilbert]第388—429页，以及赫尔曼·外尔的悼文[Weyl]。

[2] 许多读者都很熟悉，$\sqrt{2}$ 是一个无理数。（正如上一章中所说，这意味着它不可能表示为以自然数为分子分母的分数，或者说，它的小数表示是不重复的。）根据这个事实，就可以对下述定理给出一个优美的非构造性的证明：存在着无理数a和b，使得a^b 是有理数。

在证明过程中，我们用字母q表示$\sqrt{2}^{\sqrt{2}}$。现在，q必然或者是有理数，或者是无理数。如果q是有理数，那么我们就可以设

$a=b=\sqrt{2}$ 而得到需要证明的结果。如果q是无理数，那么我们可以设a=q, b=$\sqrt{2}$，然后，

$$a^b = q^{\sqrt{2}} = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{(\sqrt{2}\cdot\sqrt{2})} = (\sqrt{2})^2 = 2$$

于是，我们又一次得到，一个无理数的无理数幂次结果为一个有理数。这个证明是非构造性的，因为它并没有给出满足该定理的具体的a和b，而仅仅给出了两种可能性，其中之一必定成立。（事实上，q是无理数，但对于这个事实还没有简单的证明。）

[3]　在代数不变量理论中，所谓的幺模变换是特别令人感兴趣的。这些变换的形式是，用表达式（py+q）/（ry+s）来代换一个方程中的一个未知量（比如x），其中y是一个新的未知量，p、q、r、s是使得ps-rq=1或-1的特定的数。布尔发现对于一般的二次方程$ax^2+bx+c=0$（其中a, b，c可以代表任何数），表达式b^2-4ac（被称为方程的判别式）在这种幺模变换下是一个不变量：

在给定的二次方程中做了所说的代换并清除了分数之后，一个以y为未知数的新的二次方程就产生了。这个方程可以写成$Ay^2+By+C=0$，其中A, B，C取决于a、b、c、p、q、r、s所有这些量。b^2-4ac是不变量的意思是，新的方程的判别式与给定的方程的判别式是相同的，即$b^2-4ac=B^2-4AC$。

如果没有ps-rq=±1这个特定条件，这两个判别式的关系就是：

$$B^2-4AC=(b^2-4ac)(ps-rq)^2.$$

希望自己算出这个结果的读者可以先写下

$$ax^2+bx+c=a(x-x_1)(x-x_2),$$

其中x_1，x_2是方程的两个根，应用二次公式

$$b^2-4ac=4a^2(x_1-x_2)^2.$$

[4] 在那篇悼文中[Weyl]，赫尔曼·外尔写道：

的确，通过发现新的思想和引入新的强有力的方法，他不仅把这个问题提高到了克罗内克和戴德金为代数所设定的水平，而且还把工作做得如此彻底，以至于他就差完成它了……带着无可非议的自豪感，他在

他的论文“论完全不变量系”（Über die vollen Invariantensysteme）中以下面的话结尾：“我相信[代数]不变量理论的最重要的目标已经达到了。”随后便退到了幕后。

[5] 古典形式的数论处理的是可以在自然数1，2，3，……中发现的关系和样式，特别是有关素数和整除性的问题。在代数数论中，这些问题中有一些与某些代数方程的整数根（实根或复根）有关。高斯已经研究了形式为$m+n\sqrt{-1}$ 的数，其中m和n都是普通整数，发现了这些“高斯整数”中哪些是素数，并且证明了这样一条定理，即对于这些数来说，它可以以一种方式分解为素数，就像普通整数的情况那样。然而，如果我们研究形式为$m+n\sqrt{10}$的数，那么这个结论就不对了。一个反例是，

$$6=2\cdot 3=(2+\sqrt{10})(-2+\sqrt{10})$$

可以证明，其中2、3、$2+\sqrt{10}$和$-2+\sqrt{10}$ 都是素数，于是唯一的因式分解就不成立了。康托尔的朋友戴德金和他的强硬对手克罗内克都已经说明了如何通过考虑所谓的素数理想来重新得到唯一的因式分解。在散步的时候，希尔伯特和他的朋友胡尔维茨已经讨论了这些相互竞争的立场，并认为两者都是糟糕的。与此相反，希尔伯特在《数论报告》中的处理是优美的。

[6] [Hilbert]，第400，401页。

[7] 希尔伯特在演讲中没有时间一一讲完23个问题，他只能挑一些来讲。关于整个讲演以及由玛丽·温斯顿·纽森英译的所有23个问题，参见[Browder]，第1—34页。

[8] 参见[Browder]。（我是这篇关于第十问题的文章的合作者。）

[9] 引文详细的内容见第4章结尾。

[10] [van Heijenoort]，第129—138页。

[11] [Poincaré]，第3章。

[12] 同上。

[13] 同上。

[14] 对罗素的“精心设计的、使用起来很不方便的”分层的技术表达是“分支类型论”。

[15] 就像弗雷格的《概念文字》一样，《数学原理》中主要的推理规则是从两个形式分别为$\mathscr{A} \supset \mathscr{B}$和$\mathscr{A}$的公式推出相应的公式$\mathscr{B}$（所谓的肯定前件推理或分离规则）。尽管弗雷格非常清楚这一点，但怀特海和罗素却通过把这条规则表达为他们的“原始命题”而使人难以理解：真命题所蕴含的任何东西都为真（[White-Russ]，第94页）。

[16] [Brouwer 2].

[17] 布劳威尔的博士论文是用荷兰语写成的。英译本见[Brouwer 1]，第13—97页。

[18] [van Stigt]，第41页。

[19] 引文出自布劳威尔的博士论文（[Brouwer 1]，第96页）。

[20] 在这个非构造性证明的例子中，断言“q必定或者是有理数，或者是无理数”需要用到排中律。

[21] 外尔因康托尔和戴德金的著作中使用了所谓的非直谓定义而特别感到沮丧。如果被定义的项是该定义所使用的集合的一个成员，那么该定义就是非直谓的。从数学对象逐步被构造出来的哲学观点来看，这样一个定义是令人不快的，因为相关集合不可能先于它的一个元素被构造出来。相反的哲学观点被称为柏拉图主义，即数学对象事先就已经存在了，定义只是把它们挑选出来（比如：玛蒂尔达是屋子里最高的人），这是外尔所不能接受的。

[22] 这是1922年所做的一个讲演（首先是在哥本哈根，然后是在汉堡）的一部分。感谢瓦尔特·费尔舍使我注意到希尔伯特的激烈言辞与他那个时代之间的关系。讲演的全文（英译）可以在[Mancosu]，第198—214页中找到。我发现该译本非常准确，但没有完全表达原文中的激

情。我参考了几个译本和原文（[Hilbert]，第159—160页），以期能译得更好。

[23]　[Reid]，第137—138、144、145页。关于德国知识界人士的宣言的背景，参见[Tuchman]，第322页。

[24] [Reid]，第143页。

[25] [Hilbert]，第146页（我的翻译）。

[26]　　　　希尔伯特的纲领在一篇有趣的文章中得到了讨论，载[Mancosu]，第149—197页。基于清楚表明希尔伯特思想演化的未发表的材料所做的详细讨论和分析，参见[Sieg]。关于贝尔奈斯的贡献的有趣信息，参见[Zach]。关于冯·诺依曼把直觉主义归于荒谬，参见[Mancosu]，第168页。应当提到，尽管希尔伯特关于哪些方法可以算作“有限”的描述从来也不够清楚，但普遍认为，他的看法甚至比布劳威尔准备接受的方法更为严格。

[27] [van Heijenoort]，第373页。

[28] [van Heijenoort]，第376页。

[29] [van Heijenoort]，第336页。

[30] [Reid]，第187页。

[31] [van Stigt]，第272页。

[32] [van Stigt]，第110页。

[33] [van Stigt]，第285—294页；[Mancosu]，第275—285页。

[34] 关于计算机科学中的直觉主义逻辑，参见[Constable]。

[35] [Hilbert]，第378—387页。

[36] [Dawson]，第69页。

第6章

[1] 关于爱因斯坦谈论哥德尔投票给艾森豪威尔一事，参见[Dawson]，第209页。我很幸运能够读到这本哥德尔的传记，它写得相当好。我也利用了1983年在萨尔茨堡举行的哥德尔专题讨论会（我也有幸被邀请）的简明文集[Gödel-Symp]。在格奥尔格·克莱瑟尔所著的悼念回忆文章[Kreisel]中有很多有趣的材料，但不幸的是它们并不是完全可靠。克莱瑟尔一度曾是哥德尔的密友。在[Gödel]，第一卷，第1—36页有一篇关于哥德尔的简短而感人的传记，作者是逻辑学家所罗门·费弗尔曼。

[2] [Gödel]，第三卷，第202—259页。

[3] [Dawson]，第58、61、66页。

[4] [Gödel-Symp]，第27页。

[5] “弗雷格—罗素—希尔伯特”是一种过于简化的说法。希尔伯特所挑选的基本逻辑只是弗雷格的系统和罗素的体系的一部分，它今天被称为一阶逻辑。

[6] 关于哥德尔谈论逻辑学家们的盲目一事，参见[Dawson]，第58页。哥德尔的博士论文以及以此为基础发表的文章可以在[Gödel]第一卷，第60—123页找到（原始的德文和英译文都有）。在这些文章前有伯顿·德莱本和让·凡·海金诺特写的一个富有启发性的导论，参见第44—59页。

[7] 虽然希尔伯特的元数学中的有限性方法经常会被说成是“直觉主义的”，但希尔伯特所思考的方法很有可能要比布劳威尔所允许的严格得多。关于对该问题的讨论，参见[Mancosu]，第167—168页。

[8] [Gödel]，第一卷，第65页。

[9] 当一串0被放在一个十进位制表示的自然数之前时，该自然数并不会改变。例如：17=017=0017，以此类推。所以字符串L_y 和，L_y 和，L_y 都可用17这个数来编码。然而，由于我们对以逗号开头的字符串没有兴趣，所以这种含糊不清就无须考虑。虽然它并不真的重要，但这里

仍然可以提到，在哥德尔用来为字串编码的实际技巧中，他并没有使用十进制的数字来表示数。根据一个自然数可以被唯一地分解为素因子的乘积这一事实，他把分配给符号的代码数作为指数加在了相应的素数上。一个简单的例子将使我们看清楚其中的区别。字符串L（x, y）在我们的方案中将被编码为186079，在哥德尔的方案中，其代码数将是2^1 3^8 5^6 7^0 11^7 13^9 。

[10] 这篇划时代的文章有许多英译本，其中最好的译本（以及哥德尔所认可的一个译本）均见于[Gödel]，第一卷，第144—145页（对页有德文原文）以及[van Heijenoort]，第596—616页。对哥德尔发现他的不完备性定理的过程有兴趣的读者可以参见[Dawson]，第61页。

[11] 为了避免使用像“真理”这样的在哲学上受到怀疑的概念，哥德尔诉诸于一种技术上的替代物，他称之为ω一致性，它是一种加强了的一致性。于是，他的定理的正确表述就是：如果PM是ω一致的，那么就存在着一个命题U，U和¬ U在PM中都是不可证的。几年以后，J.B.罗塞尔说明了如何用普通的一致性假设来替代ω一致性假设，从而做出了重要的改进。与同时期的其他一些工作（特别是阿兰·图灵的工作，这将在接下来的一章中讨论）结合在一起，我们就可以以一种吸引人的形式把哥德尔的结论表述出来：无论在PM中加入什么样的附加公理，只要新的公理是通过一种算法来指定的，而且它们不会导致一个可证明的矛盾（比如一个形式为A∧¬ A的陈述），那么系统中就必定存在着一个不可判定的命题U。

[12] PM系统过于复杂，我们在这里难以详述。故此，我们用较为简单的系统PA来显示在构造不可判定命题时所牵涉的一些因素。PA可以用16个符号建立起来：

$$\supset \ \neg \ \vee \ \wedge \ \forall \ \exists \ \mathbf{1} \ \oplus \ \otimes \ x y z \ (\quad) \ ' \ \doteq$$

符号1，+，×，=之所以写成古怪的样子，是为了强调它们将仅仅被视为符号，同时也暗示它们原有的含义。字母x，y，z是表示自然数的变元，由于所需要的变元不止3个，所以我们可以把符号'加在那些字母上方，从而可以产生出任意多个变元。于是，y′和z‴都是变元。由于这里

的符号超过了10个，我们将使用一种编码表，其中每一个符号都被一对十进制数字所替代：

$\supset$	$\neg$	$\vee$	$\wedge$	$\forall$	$\exists$	$\underline{\mathbf{1}}$	$\oplus$	$\otimes$	$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{z}$	(	)	′	$\doteq$	,
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	
10	11	12	13	14	15	21	22	23	31	32	33	41	42	43	44	

自然数本身又被特定的字符串所表示，这些字符串被称为数字符号（numerals），如下所示：

数字符号	被表示的数	代码
$\underline{\mathbf{1}}$	$\underline{1}$	21
$(\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})$	2	4121222142
$((\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})\oplus\underline{\mathbf{1}})$	3	414121222142222142
$(((\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})\oplus\underline{\mathbf{1}})\oplus\underline{\mathbf{1}})$	4	41414121222142222142222142
…	…	…

大多数由这些符号所组成的字符串都是毫无意义的，例如：

$$\exists\oplus\otimes\boldsymbol{x}\forall\neg \text{ 或者 } \doteq\supset\underline{\mathbf{1}}'\ (\ \ )$$

它们的代码分别是152223311411和441021434142。但某些被称为句子的字符串可以被用来表示关于自然数的真假命题。于是，字符串

$$((\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})\otimes(\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})\doteq(((\underline{\mathbf{1}}\oplus\underline{\mathbf{1}})\oplus\underline{\mathbf{1}})\oplus\underline{\mathbf{1}}))$$

的代码是

414121222142234121222142444141412122214222214222214242，

它表示真命题2乘2等于4。而

$$((\underline{1}\oplus\underline{1})\otimes(\underline{1}\oplus\underline{1})\doteq((\underline{1}\oplus\underline{1})\oplus\underline{1})$$

则表示假命题2乘2等于3。句子

$$(\forall x)(\neg(x\doteq\underline{1})\supset(\exists y)(x\doteq(y\oplus\underline{1})))$$

的代码是

4114314241114131442142104115324241314441322221424242，

它表示除1以外的所有自然数都有前趋。

为了完成我们对PA的描述，有必要把某些特定的句子指定为公理和用于从公理推出可证句子的推理规则。从公理开始，到PA中的一个可证句子结束，其间的整个步骤被称为该句子的证明。虽然仔仔细细做完此事会把我们带得太远，但我们还是来考虑一个简单的例子：

$$(\forall x)\neg(\underline{1}\doteq(x\oplus\underline{1}))$$

它表示这样一个命题：1不是任何自然数的直接后继。这个句子完全可以被选为一条公理。既然句子开头的符号∀表示某种属性对于所有自然数都适用，那么一条自然的推理规则就是，允许用一些数字符号来替换x（除去全程量词（∀x）之后）。这只是从一个普遍陈述推出它的一个特殊情形。下面是一个简单的例子：

$$\frac{(\forall x)\neg(\underline{1} \doteq (x \oplus \underline{1}))}{\neg(\underline{1} \doteq (\underline{1} \oplus \underline{1}))}.$$

该结论是PA中的一个可证句子，它是通过用1替换变元x而得到的，它表示1和2是不等的这一事实。

除了用来表示命题的字符串外，还有其他一些被称为一元字符串，它们可以被用来定义自然数的集合。这些字符串将包含符号x，但不包含“量词”（∀x）或（∃x）（虽然它可以包含像y或x″这样的关于其他变元的量词）。此外，一元字符串有一种至关重要的性质：假如所有x都被某个数字符号所代替，那么由此得到的字符串将是一个句子。以下是一个例子：

$$(\exists y)(x \doteq ((\underline{1} \oplus \underline{1}) \otimes y))$$

它的代码是

41153242413144414121222142233324242.

假如x被数字符号$(\underline{1} \oplus \underline{1})$所代替，则我们就得到了真句子

$$(\exists y)((\underline{1} \oplus \underline{1}) \doteq ((\underline{1} \oplus \underline{1}) \otimes y))$$

假如x被数字符号1所代替，我们就得到了假句子

$$(\exists y)(\underline{1} \doteq ((\underline{1} \oplus \underline{1}) \otimes y))$$

可以认为这个一元字符串为偶数集提供了一种定义。以下是一个更为复杂的一元字符串：

$$(\forall y)(\forall z)((x \doteq (y \otimes z)) \supset ((y \doteq 1) \vee (y \doteq x)))$$

它的代码是

41143242411433424241314441322333424210414132441421241324443

它定义了由1和全部质数所组成的集合。给定一个一元字符串A和自然数n，我们将用符号[A：n]来表示用表示数字n的数字符号来替换A中的x所得到的句子。例如：

$$[\ (\exists y)(\ x \doteq ((\ \underline{1} \oplus \ \underline{1}) \otimes y)) : 2\ ]$$

表示句子

$$(\exists y)((\ \underline{1} \oplus \ \underline{1}) \doteq ((\ \underline{1} \oplus \ \underline{1}) \otimes y))$$

现在，我们就可以解释如何用哥德尔的方法生成PA中的一个句子U，该句子表示U本身在PA中是不可证的这一命题。使用被分配给一元字符串的码数，我们就可以依照它们的代码大小将其全部排列起来。在这种排序中，代码最小的一元字符串是（x=1），即使连它的代码也是4131442142，超过了40亿。我们用A_1 来表示这个一元字符串，并想象所有的一元字符串都依照它们代码的大小排成一个序列：

A_1，A_2，A_3，……

因为这些都是一元字符串，所以对于任意的自然数n, m，字符串[A_n：m]就将是一个句子。这些句子中的某一些在PA中将是可证的，另一些则是不可证的。对于每一个n，我们都可以考虑由那些使得字符串[A_n：m]在PA中是不可证的m所组成的集合。回忆一下我们对康托尔的对角线方法所做的讨论，我们看到这样一个集合可以被看成一个以n为

标签的包裹。运用对角线方法，也就是说，把包裹中的一个元素当作标签，我们就可以构造出一个由那些使得[A_n ：m]在PA中不可证的n所组成的集合K。利用PA中的可证明性可以在PA中进行定义这一事实，我们就可以找到一个一元字符串B，它在PA中定义了集合K。现在，必然存在着某个数q使得B=A_q ，因为所有的一元字符串都包含在A_s 的序列中。于是，对于每一个自然数n，句子[A_q：n]都表示命题

[A_n：n]在PA中是不可证的。

特别地，当n被赋予q值时，我们可以看到，[A_q：q]表示命题

[A_q：q]在PA中是不可证的。

因此，[A_q ：q]是PA中的一个句子，它表示它本身在PA中不可证这一命题。

[13] 在哥德尔证明了PM的一致性不能用PM所包含的全部数学资源进行证明之后，我们很自然地会得出结论说，用希尔伯特将会允许的有限性方法来证明这种一致性是不可能成功的。这肯定也是冯·诺依曼的结论。哥德尔本人并不是如此确信，他仍然抱有一种希望：也许存在着某些在PM内部不被允许的方法可以被认为是有限性方法，它们将导致对一致性的证明。在哥德尔的发现做出之后的几十年中，有些人声称自己提出了某种方法，它能够满足这种标准。结果，尽管几乎没有人会宣称，已获证明的一致性定理给相关系统的有效性增加了任何信心，但希尔伯特的证明论作为一个研究领域仍然持续不断地蓬勃发展。

[14] 那些主要被用于软件业的程序设计语言（比如C语言和FORTRAN语言）通常被说成是命令式的，这是因为可以认为用这些语言编写的一行行程序就是计算机所执行的命令。像C++这样的模件式语言也是命令式的。在那些所谓的功能程序设计语言（比如LISP语言）中，程序的字符行就是操作的定义。它们不是告诉计算机做什么，而是定义计算机将要提供什么。哥德尔的特殊语言非常像一种功能程序设计语言。

[15] 回到我们所建议的那种特殊编码方式的PA的例子，我们可以考察在把元数学概念翻译成数值运算的过程中所涉及的一些问题。我们可

以提出的第一个问题是，给定某一字符串的代码数，我们如何能够说出该字符串的长度？现在，由于我们用两个阿拉伯数字来表示一个符号，所以答案很简单：字符串的长度等于代码中的数字数目的一半。对于一个代码数r，我们将把相应的字符串的长度记为 $\mathscr{L}(r)$ 。给定两个字符串，我们总是可以通过把第二个字符串放到第一个字符串后面来构成一个新的字符串。第二个问题是，如果给定的两个字符串的代码分别是r和s，那么所组成的新的字符串的代码是什么？答案可以通过公式 $r10^{2\mathscr{L}(s)}+s$ 来计算。这是因为，r乘以10的 $2\mathscr{L}(s)$ 次幂可以表示在r后加了多少个0，这也就是s中的数字的个数。遵循哥德尔的方法，我们把它写作r*s。现在假定r和s是两个句子的代码，如果我们在这两个句子之间插入符号⊃，并将结果用括号括起，那么这样得到的新句子的代码是什么？通过查译码表，我们知道答案是41*r*10*s*42。按照这种方法不断进行下去，即使是更复杂的元数学概念也可以被翻译成算术运算。

[16] 中国剩余定理显然可以追溯到公元11世纪时的中国。该定理可以用如下的练习加以说明：试找到一个数，它被6除余2，被11除余5。经过一番简单的试验之后，可知这个数是38。中国剩余定理保证我们总可以找到一个数，它被给定的数所除后余数是另一个给定的数，只要这些给定的除数两两之间没有公因数（当然不包括1）。例如，肯定存在着这样一个数，当它被69、17、25、91所除后，余数分别为13、7、10、11。但是假如我们把7换成14，那么这个结论就不能保证成立了（因为10和14有公因数2）。哥德尔把中国剩余定理用作一个编码装置：设计一组两两之间没有公因数的除数，然后让它们都除以同一个数，我们就可以定出一连串数字。由于“余数”在算术的基本语言中是容易定义的，所以它可以被用来表示在该种语言中包含自然数序列的关系。

哥德尔用中国剩余定理去为有限的自然数序列编码的技巧在我个人的职业生涯中扮演着重要的角色。作为我博士论文研究（1950年为普林斯顿大学所接受）的一部分，我钻研了希尔伯特的第十问题，而中国剩余定理对于我所能够获得的部分结果相当重要。后来我与希拉里·普特南以及朱莉娅·鲁宾逊共同开展的工作仍然需要用到此定理。解决希尔伯特的第十问题的关键的最后一步是俄国22岁的数学家尤里·马蒂雅谢维奇在1970年做出的。有兴趣的读者可以参考专门为普通读者所写的文章[Dav-Hersh 2]。

[17]　卡尔纳普、海丁和冯·诺依曼在柯尼斯堡所作演讲的全文可以在[Ben-Put]，第41—65页找到。

[18] 关于哥德尔在柯尼斯堡圆桌讨论会上的评论的完整陈述（德文原文和英译文）以及约翰·道森启发式的评论，参见[Gödel]，第一卷，第196—203页。也参见[Dawson]，第68—71页。

[19] [Dawson]，第70页。

[20] [Goldstine]，第174页。

[21] 同上。

[22] 这项研究包含了非常大的超限基数，它超出了本书的范围。对此曾有一篇有趣的文章，其作者是一位一流的怀疑论者，参见[Feferman]。

[23] [Dawson]，第32—33页，第277页。

[24] [Dawson]，第34页。

[25] [Dawson]，第111页。

[26] [Gödel-Symp]，第27页。

[27] 这些贡献中最有趣的部分与布劳威尔的学生海丁所发展的某些形式系统有关，海丁试图将布劳威尔的基本思想全都包含在内。布劳威尔本人确信，没有一种被精确定义的形式语言能够恰当地处理他的概念，但他还是对海丁的工作勉强表示了兴趣。海丁的系统中有一个是HA（表示海丁算术），它与PA非常类似，只是背后的逻辑规则是布劳威尔认为能够接受的规则，而不是弗雷格的规则。特别是，排中律在HA中是不成立的。哥德尔找到了一种简单的方法可以将PA翻译成HA，于是，与直觉主义比古典数学狭窄的想法相反，这里直觉主义在某种意义上包含了古典数学。特别是，任何对HA的一致性的证明立即可以被翻译成对PA的一致性的证明。

[28] [Dawson]，第103—106页。

[29] [Gödel-Symp]，第20页。

[30] [Dawson]，第142页，第146页。

[31] [Dawson]，第91页。

[32] [Dawson]，第147页。

[33] [Dawson]，第143—145页，第148—151页。

[34] [Dawson]，第153页。

[35] [Browder]，第8页。

[36] 更精确地说，哥德尔说明的是：如果像PA这样的系统或者那些建立在集合论公理上的系统是一致的，那么即使是把连续统假设作为一条新的公理加进去，它也仍然是一致的。因此，如果这些系统是一致的，那么在其内部连续统假设就不可能被否证。

[37] 这场斗争很是激烈。杰出的逻辑学家所罗门·费弗尔曼在一篇发表于本书写作期间的文章[Feferman]中称，连续统假设“本来就是含糊不清的”。在经历了最初的一些摇摆不定之后，哥德尔终于认为连续统假设绝不是含糊不清的，事实上，它是一个完全有意义的断言，并且很可能是错的。逻辑学家W.休·伍丁近期的工作强烈地暗示哥德尔是正确的。

[38] [Gödel]，第二卷，第108，186页。

[39] [Gödel]，第三卷，第49—50页。

[40] [Gödel]，第二卷，第140—141页。

[41] [Gödel]，第三卷收录了大部分以前未曾发表的哥德尔的著作。

[42] [Dauben 2]，关于哥德尔希望让鲁宾逊做他的继任者，参见第458页；关于所引用的信件，参见第485—486页。

第7章

[1] [Huskey]，第300页。

[2] [Ceruzzi]，第43页。

[3] 我有幸看到了安德鲁·霍奇斯关于图灵的令人辛酸的文笔优美的传记；参见[Hodges]。

[4] [Hodges]，第29页。

[5] 图灵用生动的话语表达了他对自己死去的朋友的感情：阿兰“崇拜他所走过的地面”，“他使每个人都显得如此平凡”。参见[Hodges]，第35，53页。

[6] [Hodges]，第57页。

[7] [Hodges]，第94页。

[8] 事实上，希尔伯特并不是这样提出判定问题的：他想知道的是，给定一个一阶逻辑表达式，如何有一种程序来判定它是否在任何可能的解释中都是有效的。不过，在哥德尔证明了他的完备性定理之后，情况就已经很清楚，这里问题被陈述的形式等价于希尔伯特的表述。

[9] 关于判定问题的工作主要是处理被称为前束式的表达式。它们是包含$\neg$，$\supset$，$\wedge$，$\vee$，$\exists$，$\forall$等逻辑符号的表达式，其中所有的存在量词（$\exists$……）和全程量词（$\forall$……）都先于所有其他符号出现在表达式的开头。不难证明，判定问题可以被还原为为这样一个问题找到一种算法，即给定一个前束式，判定它是否是可满足的，也就是说，是否存在着某种解释前束式中非逻辑符号的方法，使得它可以表示一个真句子。为了说明这个概念，考虑以下两个前束式：

$$(\forall x)(\exists y)(r(x)\supset s(x,y))$$和

$$(\forall x)(\exists y)(q(x)\wedge\neg q(y))$$。

第一个前束式是可满足的：例如，我们可以令变量x， y表示在某一时刻活着的人，r（x）表示“x是一个一夫一妻制的结婚的男人”，s（x，y）表示“y是x的妻子”；于是，根据这一解释，第一个前束式说的就是“每一个一夫一妻制的结婚的男人都有一个妻子”——这当然是一个真陈述。而第二个前束式则是不可满足的，因为无论如何选择个体域，无论如何解释符号q，这一前束式既规定所有个体都具有q所表示的性质，又规定有些个体没有这种性质。

前束式可以通过它们开头的存在量词和全程量词的特殊样式来分类。例如，前置集∀∃∀的意思是所有以（∀……）（∃……）（∀……）开始的前束式所组成的集合。在哥德尔于1932年发表的一篇论文中，他提出了一种可以检验任何属于前置集∀∀∃……∃的前束式的算法。在一年以后发表的一篇论文中，他证明了要想解决判定问题，只要提出一种算法能够检验所有属于前置集∀∀∀∃……的前束式的可满足性就足够了。这样，已经得到的结果和所需的结果之间的差距就只剩下一个全程量词∀了。

哥德尔的相关论文（德文原文和英译文）见于[Gödel]，第一卷，第230—235，306—327页。沃伦·格尔德法伯为这一卷写的说明性的导言介绍了关于这个问题的某些早期工作。

[10] [Hodges]，第93页。

[11] 图灵对这一点的讨论是更为细致的；参见[Turing 2]，第250—251页。也参见选集[Davis 1]，第136—137页。

[12] 尽管判定问题的不可解性可以用这种方法来证明，但它是相当麻烦的，因为我们需要改进图灵机的构造来处理十进制整数。为了理解图灵的实际工作，我们先来说明，判定一台带子完全空白的图灵机是否会停机的问题是不可解的。因为如果假定这个问题有一种算法，那么为了检验代码数n是否属于D，我们先写出构成代码数为n的图灵机T的五元组，然后写出使n写在图灵机带子上的五元组。把这些五元组加在T的五元组上，我们就得到了一台新的机器，它将先把n写在带子上，然后按照输入n时T的响应来运转。这台从空白带子开始的新机器最终将停机，当且仅当T从带子上的n开始时会最终停机，而后一结论成立，当且仅当n不属于D。因此，所假定的用于检验一个从空白带子开始的给定

的图灵机最终是否会停机的算法，可以被用来解决判定一个数是否属于D的不可解问题。

我们还注意到，一台给定的图灵机是否曾经印下一个符号的问题也是不可解的。这是因为，很容易使得一台图灵机无论什么时候停机，它都处于不再启用五元组的状态F。我们选择这台机器的任何五元组中都没有出现的一个新符号X，然后加入五元组：

F a：X*F

其中a可以是出现在原始五元组中的任何符号。无论最初的机器什么时候停机，这台新机器都将印上X。于是我们便得到，没有算法可以判定一台从空白带子开始的图灵机是否将印上某一符号。这是图灵用一阶逻辑语言表达的问题，于是我们便得到了判定问题的不可解性。

[13] [Turing 2]，第129—132页。

[14] [Davis 2].

[15] 关于图灵博士论文的重印本，参见[Davis 1]，第155—222页。我们所提到的分层结构还拓展到了康托尔的超限数，所以在第一个、第二个、第三个……系统之后的系统数将是ω，然后是ω+1等。

[16] [Hodges]，第131页。

[17] [Hodges]，第124页。

[18] [Hodges]，第145页。对于那些熟悉柯尔莫果洛夫和柴廷关于复杂性描述的后期工作的读者来说，这个游戏可以很好地说明冯·诺依曼正在沿着那些思路进行思考。

[19] [Hodges]，第545页。

[20] 图灵在布莱奇利庄园的合作者之一——数学家彼得·希尔顿叙述了图灵在地方军的冒险经历。参见[Hodges]，第232页。

[21] 完成这项工作绝非一人之功。也许贡献最大的人是W.T.塔特。关于塔特教授对这个问题的技术描述，以及图灵所起的作用，参见网址

http：//home.cern.ch/frode/crypto/tutte.html。

第8章

[1] 这里的引文参见[Goldstine]，第22页。关于爱达·洛甫莱斯的那本迷人的传记[Stein]暗示，大部分有关她的事情皆为虚构而非事实。

[2] [Goldstine]，第120页。

[3] 阿塔纳索夫的机器是为解决联立的线性方程组而设计的。这种问题的一个例子是

$2x+3y-4z=5$

$3x-4y+2z=2$

$x-3y-5z=4$

机器最多可以处理含有30个未知量的30个方程。

[4] [Lee]，第44页。本章的传记材料主要来自于这本书。

[5] [Burks-Burks].

[6] 微分分析器含有若干旨在计算定积分的数值近似的部件。ENIAC也含有可以完成同样工作的部件，但由于使用了著名的算法，所以结果更为准确。

[7] [Goldstine]，第186，188页。

[8] 虽然冯·诺依曼的“关于EDVAC的报告草案”流传很广，非常有影响，然而只是到了1981年，它才作为一本书的附录发表，而这本书对他的贡献的重要性是表示怀疑的。

[9] [McCull-Pitts]；[von Neumann 2]，第319页。

[10] [Goldstine]，第191页。

[11] [Randell]，第384页。

[12] [Goldstine]，第209页；[Knuth]。

[13] [von Neumann 2]，第1—32页。

[14] [von Neumann 2]，第34—97页。

[15] 关于极力抹杀冯·诺依曼对计算机的贡献，并且完全忽略图灵的贡献的研究著作的例子，参见[Metrop-Worlt]和[Stern]。关于埃克特备忘录（这是一位工程师的“揭发”）中的节选，参见[Stern]，第28页。

[16] [Stern]讨论了埃克特—莫齐利在商业上的努力的兴衰过程。

[17] 这里引用的对ACE报告的分析是一篇优秀的论文[Carp-Doran]。报告本身可以在[Turing 1]，第1—105页中找到。在许多年里，它只是以油印形式流传，并不容易看到。

[18] 用现代的术语来说，图灵所建议的是用堆栈来代替子程序处理。堆栈可以以后进先出（LIFO）的结构对数据进行排列。于是，当一次计算被中断，以调用一个预先编好的子程序时，它必须记住子程序结束时所要回到的位置。由于子程序可以调用其他子程序，这将导致一个记忆这种中断位置的堆栈。图灵建议把放入堆栈的操作称为“bury”，从堆栈的“顶部”释放称为“unbury”。（今天用的是术语是“push”和“pop”。）

[19] [Hodges]，第352页。

[20] [Turing 1]，第106—107页。

[21] [Turing 1]，第102—105页；[Hodges]，第361页。

[22] [Metrop-Worlt]；[Stern].

[23] [Goldstine]，第191—192页。

[24] [Turing 1]，第25页。

[25] [Davis 3].

[26] [Whitemore].

[27] [Marcus]，第183—184页。所引的书是恩格斯著名的《英国工人阶级状况》。

[28] [Lavington]，第31—47页。

[29] [Goldstine]，第218页。

[30] [Hodges]，第149页。

第9章

[1] [Turing 1]，第122页。

[2] 在美国科学促进会上发言的5位计算机科学家以及他们讲演的题目分别为：Joseph Y.Halpern，*Epistemic Logic in Multi-Agent Systems*；Phokion G.Kolaitis，*Logic in Computer Science-An Overview*；Christos Papadimitriou，*Complexity As Metaphor*；*Moshe* Y.Vardi, From Boole to the Pentium；Victor D.Vianu，*Logic As a Query Language.*

[3] 关于约瑟夫·外岑鲍姆的简短的传记注释，参见[Lee]，第724页。

[4] [Turing 1]，第133—160页。

[5] 这篇文章[Searle]包含了他关于这个主题所写的其他一些东西。这段话实际上是对雷·库茨维尔的一本流行读物的评论。我并没有帮助库茨维尔反击塞尔的意思，而只是把这篇评论作为塞尔那些经常说起的观点的一个方便来源。

[6] 在算法过程的概念已经被图灵、丘奇和其他人阐明之后，哥德尔定理只能以这种方式进行表述了。

[7] 彭罗斯首先在他那本雅俗共赏的书[Penrose 1]中做了这个断言。尽管有一些逻辑学家已经试图纠正他的错误，但他仍然坚持自己误入歧

途的观点。我曾就这个话题写过一篇文章，参见[Davis 4]。[Penrose 2]中包含了对他的批评者的回应，[Davis]是我对他的回答所做的回复。

[8] 更多的相关信息和参考书，参见[Gödel]，第二卷，第297页。

参考书目

[Aiton]Aiton, E. J.，*Leibniz*：*ABiography*. Bristol, UK, and Boston, MA：Adam Hilger Ltd.，1985.

[Baker-Hacker]Baker, G. P.，and P.M.S.Hacker.*Frege*：*LogicalExcavations*. New York：Oxford University Press and Oxford：Basil Blackwell，1984.

[Barret-Ducrocq]Barret-Ducrocq, F.*Love in the Time of Victoria*. New York：Penguin Books. 1992.

[Ben-Put]Benacerraf, P. and H.Putnam.*Philosophy of Mathematics*；*Selected Readings*， 2e.London and New York：Cambridge University Press，1983.

[Boole 1]Boole, G.*The Mathematical Analysis of Logic, Being an Essay towards a Calculus of Deductive Reasoning*. Cambridge, UK：Macmillan, Barclay, and Macmillan，1847.

[Boole 2]——.*An Investigation of the Laws of Thought on which Are Founded the Mathematical Theories of Logic and Probabilities*. London：Walton and Maberly，1854；reprinted, New York：Dover，1958.

[Boole 3]——.*A Treatise on Diferential Equations*. 5e. Cambridge, UK：Macmillan，1865.

[Boolos]Boolos, G.“Frege's Theorem and the Peano Postulates，”*TheBulletinofSymbolicLogic*，vol. 1（1995），pp.317—326.

[Bourbaki]Bourbaki, N，*Eléments d’Histoire des Mathématiques*，2e.Paris：Hermann，1969.

[Britannica]*TheEncyclopaediaBritannica*，11e. Cambridge, UK：

Cambridge University Press，1910，1911.

[Brouwer 1]Brouwer, L. E.J.，*Collected Works* ，vol.I.Amsterdam：North-Holland，1975.

[Brouwer 2]——.“Life, Art, and Mysticism，”*NotreDaweJournalofFormalLogic* ，vol. 37（1996），pp.389—429.

Translated by W. P.van Stigt.See also the introduction by the translator, pp.381—387.-

[Browder]Browder, F.，de.“Mathematical Developments Arising from Hilbert's Problems.”*Proceedings of Symposia on Pure Mathematics.* vol.ⅩⅩⅧ.Providence：American Mathematical Society，1976.

[Burks-Burks]Burks, A. W.，and A.R.Burks.“The ENIAC：First General-Purpose Electronic Computer，”*Annals of the History of Computing*， vol.3（1981），pp.310—399.

[Bynum]Bynum, T. W.，ed.and transl.*ConceptualNotationandRelatedArticles* by G.Frege.London and New York：Oxford University Press，1972.

[Cantor 1]Cantor，*G. Gesammelte Abhandlungen.* Berlin：Julius Springer，1932.

[Cantor 2]——.*ContributionstotheFoundingoftheTheoryofTransfniteNumbers.* La Salle, IL：Open Court，1941. Translated from the German with an introduction and notes by P.E.B.Jourdain.

[Carp-Doran]Carpenter, B. E.，and R.W.Doran.“The Other Turing Machine，”*Computer Journal.* vol.20（1977），pp.269—279.

[Ceruzzi]Ceruzzi, P. E.Reckoners, the*PrehistoryoftheDigitalComputer*, *fromRelaystotheStoredProgram Concept*， 1933—1945.Westport, CT：

Greenwood Press，1983.

[Constable]Constable, R. L.et al.*Implementing Mathematics with the Nuprl Proof Development System.* Englewood Cliffs, NJ：Prentice-Hall.1986.

[Couturat]Couturat, L.*LaLogiquedeLeibnizd'AprèsdesDocumentsInédits.* Paris：F.Alcan，1901.Reprinted Hildesheim：Georg Olms，1961.

[Craig]Craig, G. A.*Germany 1866—1945.* London and New York：Oxford University Press，1978.-

[Daly]Daly, D. C.“The Leaf that Launched a Thousand Ships，”*NaturalHistory*，vol.105，no.1（Jan.1996），pp.24—32.

[Dauben 1]Dauben, J. W.*Georg Cantor*：*His Mathematics and Philosophy of the Infnite.Princeton*， NJ：Princeton University Press.1979.

[Dauben 2]——.*Abraham Robinson*：*The Creation of Nonstandard Analysis, a Personal and Mathematical Odyssey.* Princeton, NJ：Princeton University Press. 1995.

[Davis 1]Davis, M.，ed.*The Undecidable.* Hewlett, NY：Raven Press，1965.

[Davis 2]——.“Why Gödel Didn’t Have Church’s Thesis，”*InformationandControl*，vol.54（1982），pp.3—24.

[Davis 3]——.“Mathematical Logic and the Origin of Modern Computers，”*Studies in the History of Mathematics.* Washington, DC：Mathematical Association of America，1987，pp. 137—165.Reprinted in R.Herkerr, ed.*TheUniversalTuringMachine—AHalf-CenturySurvey.* Hamburg and Berlin：Verlag Kemmereer&Unverzagt，1988；and London and New York：Oxford University Press，1988.

[Davis 4]——.“Is Mathematical Insight Algorithmic?”*BehavioralandBrainSciences*，vol. 13（1990），pp.659—

660.

[Davis 5]——.“How Subtle is Gödel’s Theorem?More on Roger Penrose，”*Behavioral and Brain Sciences*， vol.16（1993），pp.611—612.

[Dav-Hersh 1]Davis, M. and R.Hersh.“Nonstandard Analysis，”*ScientificAmerican*，vol.226（1972），pp.78—86.

[Dav-Hersh 2]——.“Hilbert's 10th Problem，”*Scientifc American* ，vol. 229（1973），pp.84—91.

[Dav-Sig-Wey]Davis, M.，R. Sigal, and E.Weyuker.*Computability, Complexity, andLanguages*，2e.New York：Academic Press，1994.

[Dawson]Dawson, J. W.，Jr.*LogicalDilemmas*：*TheLifeandWorkofKurtGödel.* Wellesley, MA：A K Peters，1997.

[Dummett]Dummett, M.，Frege：PhilosophyofLanguage，2e. Cambridge, MA：Harvard University Press，1981.

[Edwards]Edwards C. H, Jr.*TheHistoricalDevelopmentoftheCalculus.* New York：Springer Verlag，1979.

[Euclid]Heath, Sir T. L.，transl.Euclid's Elements（with Introduction and Commentary），vol.I.New York：Dover Publications，1956.

[Feferman]Feferman, S.，“Does Mathematics Need New Axioms?”*American Mathematical Monthly*，vol. 106（1999），pp.99—111.

[Frege 1]Frege, G.，*Wissenschaftlicher Briefwechsel.* Hamburg：Felix Meiner，1976.

[Frege 2]——.“Diary for 1924，”*Inquiry, vol.* 39（1996），pp. 303—342.Translated by R.L.Mendelsohn.

[Frege 3]——.“Über Sinn und Bedeutung，”*Zeitschrift für Philosophie*

und philosophische Kritik， new series.vol.100（1892），pp.25—50.English translation in Geach, P.，and M.Black, eds.*Translationsfromthe Philosophical Writings of Gottlob Frege.* Oxford：Blackwell.1952.

[Frege 4]——.“Rezension von：Georg Cantor. Zum Leher vom Transfiniten，”*ZeitschriftfrPhilosophieund philosophische Kritik, new series.* vol.100（1982），pp.269—272.

[Geiss]Geiss, I.，ed.*July1914*：*TheOutbreakoftheFirstWorldWar, SelectedDocuments.* New York：Charles Scribner，1967.

[Gerhardt]Gerhardt, C. I.，ed.*DiePhilosophischenSchriftenvonG.W.Leibniz.* 7 vols.Hildesheim：Georg Olms Verlagsbuchhandlung，1978.（Photographic reprint of the original 1875-90 edition.）

[Gödel]Gödel, K.*CollectedWorks.* London and New York：Oxford University Press.vol.Ⅰ，1986；vol.Ⅱ，1990；vol.Ⅲ，1995.

[Gödel-Symp]Weingartner, P.and L.Schmetterer, eds.*GödelRemembered.* Naples：Bibliopolis，1983.

[Goldstine]Goldstine, H. H.*TheComputerfromPascaltovonNeumann.* Princeton, NJ：Princeton University Press，1972.

[Grattan-Guinness]Grattan-Guinness, I.“Towards a Biography of Georg Cantor，”*AnnalsofScience*， vol. 27（1971），pp.345—391.

[Hilb-Acker]Hilbert, D.，and W. Ackermann，*GrundzgederTheoretischenLogik.* Berlin：Julius Springer，1928.

[Hilbert]Hilbert, D.*GesammelteAbhandlungen, Band* Ⅲ.Berlin and Heidelberg：Springer-Verlag，1935，1970.

[Hinsley]Hinsley, F. H.and A.Stripp, eds.*Codebreakers*：*TheInsideStoryofBletchleyPark.* Oxford and New York：Oxford University

Press.1993.

[Hodges]Hodges，*A. Alan Turing*：*The Enigma.* New York：Simon and Schusterr，1983.

[Hofmann]Hofmann. J.E.，*Leibniz in Paris 1672～1676.* London：Cambridge University Press，1974.

[Huber]Huber, K.*Leibniz.* Munich：Verlag von R. Oldenbourg，1951.

[Huskey]Huskey, V. R.and H.D.Huskey，“Lady Lovelace and Charles Babbage，”*AnnalsoftheHistoryof Computing*， vol.2（1980），pp.299—329.

[Kagan]Kagan, D.*On the Origins of War and the Preservation of Peace.* New York：Doubleday，1995.

[Kinealy]Kinealy, C.“How Politics Fed the Famine，”*NaturalHistory*，vol. 105，No.1（Jan.1996），pp.330—335.

[Kluge]Kluge, E. H.W.“Frege, Leibniz, et alia”*Studia Leibnitiana.* vol.Ⅸ（1977），pp.266—274.

[Knuth]Knuth, D. E.“Von Neumann's First Computer Program，”*ComputerSurveys*，vol.2（1970），pp.247～260.

[Kreisel]Kreisel, G.“Kurt Gödel：1906～1978，”*BiographicalMemoirsofFellowsoftheRoyalSociety*，vol.26（1980），pp.149—224；corrigenda, vol.27（1981），p.68.

[Lavington]Lavington, S.*Early British Computers.* Bedford, MA：Digital Press，1980.

[Lee]Lee, J. A.N.*ComputerPioneers.* Los Alamitos, California：IEEE Computer Society Press，1995.

[Leibniz 1]Leibniz, G. W.“Dissertatio de Arte Combinatoria，”*G.W.Leibniz*：*MathematischeSchriften, Band V.*

Hildesheim：Georg Olms Verlagsbuchhandlung，1962，pp.8—79.（Photographic reprint of the original 1858 edition.）

[Leibniz 2]——.“Letter from Leibniz to Galloys. December 1678，”*G.W.Leibniz*：*MathematischeSchriften, BandI.* Hildesheim：Georg Olms Verlagsbuchhandlung，1962，pp.182—188.（Photographic reprint of the original 1849 edition.）

[Leibniz 3]——.“Machina arithmetica in qua non additio tantum et subtractio set et multiplicato nullo, divisio vero paene nullo animi labore peragantur，”1685. English translation by M.Kormes in Smith, D.E.*A Source Book in Mathematics.* New York：McGraw-Hill，1929，pp.173—181.

[Lewis]Lewis, C. I.A.*SurveyofSymbolicLogic.* New York：Dover，1960.（Corrected version of ChaptersⅠ-Ⅳof the original edition, Berkeley：University of California Press，1918.）

[MacHale]MacHale, D.*George Boole*：*His Life and Work.* Dublin：Boole Press，1985.

[Mancous]Mancosu, P.*From Brouwerto Hilbert.* London and New York：Oxford University Press，1998.

[Mates]Mates, B.*The Philosophy of Leibniz*；*Metaphysics&Language.* London and New York：Oxford University Press，1986.

[McCull-Pitts]McCulloch, W. S.，and W.Pitts.“A Logical Calculus of the Ideas Immanent in Nervous Activity，”*BulletinofMathematicalBiophysics*，vol.5（1943），115—133.Reprinted in McCulloch, W.S.Embodiments of Mind.Cambridge, MA：MIT Press，1965，pp.19—39.

[Marcus]Marcuss, S.*Engels, Manchester, and the Working Class.* New York：W. W.Norton，1974.

[Meschkowski]Meschkowski, H.*Georg Cantor*：*Leben, Werk und Wirkung.* Mannheim：Bibliographisches Institut，1983.

[Metrop-Worlt]Metropolis, N.，and J. Worlton，“A Trilogy of Errors in the History of Computing，”*Annalsof the History of Computing* ，vol.2（1980），pp.49—59.

[Parkinson]Parkinson, G. H.R.*Leibniz—LogicalPapers.* London and New York：Oxford University Press，1966.

[Penrose 1]Penrose, R.*TheEmperor'sNewMind.* London and New York：Oxford University Press，1989.

[Penrose 2]——.“The Nonalgorithmic Mind，”*BehavioralandBrainSciences*，vol. 13（1990），pp.692—705.

[Poincaré]Poincaré，H，*Science and Method.* New York：Dover，1952.

[Purkert-Ilgauds]Purkert, W. and H.J.Ilgauds.*GeorgCantor*；*1845—1918*，Vita mathematica, vol.1.Stuttgart：Birkhauser.1987.

[Randell]Randell, B. ed.*TheOriginsofDigitalComputers, SelectedPapers* 3e.New York：Springer-Verlag，1982.

[Reid]Reid, C.*Hilbert-Courant.* New York：Springer-Verlag，1986. Originally Published By Springer-Verlag as two separate works：Hilbert，1970 and*Courant in Göttingen and New York*：*The Story of an Improbable Mathematician*，1976.

[Rucker]Rucker, R.*InfnityandtheMind*：*TheScienceandPhilosophyotheInfnite.* Boston：Birkhuser，1982.

[Searle]Searle, J. R.“I Married a Computer，”*TheNewYorkReviewofBooks*，April 8，1999，pp.34—38.

[Sieg]Sieg, W.“Hilbert's Programs：1917—1922，”*Bulletin of the Association for Symbolic Logic*， vol. 5（1999），pp.1—44.

[Siek-Wright]Siekmann, J. and G.Wrightson,

eds.*AutomationofReasoning, vol.* 1.New York：Springer-Verlag，1983.

[Sluga]Sluga, H.*Heidegger's Crisis*：*Philosophy and Politics in Nazi Germany*. Cambridge, MA：Harvard University Press，1993.

[Stein]Stein, D.*Ada*：*A Life and a Legacy*. Cambridge, MA：MIT Press 1987.

[Stern]Stern, N.*FromEniactoUnivac*：*AnAppraisaloftheEckert-MauchlyMachines*. Bedford, MA：Digital Press，1981.

[Swoyer]Swoyer, C.“Leibniz's Calculus of Real Addition，”*StudiaLeibnitiana*， vol.ⅩⅩ Ⅵ（1994），pp.1—30.

[Tuchman]Tuchman, B. W.*The Guns of August*. New York：Macmillan，1962，1988.

[Turing 1]Turing, A.*Collected Works*：*Mechanical Intelligance.* Amsterdam：North-Holland，1992.

[Turing 2]——.“On Computable Numbers with an Application to the Entscheidungsproblem，”*ProceedingsoftheLondonMathematicalSociety*，ser. 2，vol.42（1936），pp.230—267.Correction：vol.43（1937），pp.544—546.Reprinted in[Davis 1]pp.116—154.

[van Heijenoort]van Heijenoort, J.*FromFregetoGödel.* Cambridge, MA：Harvard University Press，1967.

[van Stigt]van Stigt, W. P.，*Brouwer's Intuitionism.* Amsterdam：North-Holland，1990.

[von Newmann 1]von Neumann, J.，*First Draft of a Report on the EDVAC* ，Moore School of Electrical Engineering, University of Pennsylvania，1945. First printed in[Stern]，pp.177—246.

[von Neumann 2]——.*Collected Works, vol.* 5. New York：Pergamon Press，1963.

[Welchman]Welchman, G.，*The Hut Six Story*. New York：McGramHill，1982.

[Weyl]Weyl, H.“David Hilbert and His Mathematical Work，”*BulletinoftheAmericanMathematicalSociety*, *vol*. 50（1944），pp. 612—654.

[White-Russ]Whitehead, A. N.，and B.Russell.*PrincipiaMathematica*，vol.I，2e.London and New York：Cambridge University Press，1925.

[Whitemore]Whitemore, H.*Breaking the Code.* London：Samuel French Ltd，1988.

[Zach]Zach, R.“Completeness before Post：Bernays, Hilbert, and the Development of Propositional Logic，”*Bulletin of Symbolic Logic*，vol. 5（1999），pp.331—336.

译后记

计算机技术无疑是当今最热门、应用最广的技术之一，它的作用和威力可以说无人不知、无人不晓。然而，尽管计算机技术的发展日新月异，学习它的人数也与日俱增，但很少有人了解其背后的思想，很少有人知道如此复杂精妙的机器所依据的是什么。它被设计得越来越方便合用，以至于我们几乎不再对它背后的奥秘感到好奇，计算机对大多数人来说仍然是一种神奇的东西，是一个谜。

本书作者马丁·戴维斯是纽约大学库朗数学科学研究所的名誉教授，目前在加州大学伯克利分校做访问学者。他是计算机科学发展史上的先驱人物，也是世界著名的数理逻辑学家，曾对希尔伯特第十问题有过深入的研究。他的《可计算性与不可解性》一书被誉为“计算机科学领域极少数真正的经典著作之一”。而摆在读者面前的这本书出版后即广受好评，被誉为从逻辑角度讲述计算机发展史的最好的通俗读本。粗略地说，计算机技术可以分为两大部分，其一是它的工程实现方面，另一则是它的思想或逻辑方面，但已往的关于计算机发展史或计算史的书大都只重视前者而忽视后者。当我们随手翻开一本这样的书时，眼前出现的不外乎巴贝奇、阿塔纳索夫、艾肯、埃克特、莫齐利等一串工程师的名字，讲述的是电子管、晶体管、存储芯片等的发展。而本书的主要人物却是莱布尼茨、布尔、康托尔、希尔伯特、哥德尔和图灵等人，这是怎么回事呢？原来，戴维斯认为通常所讲的历史只是故事的一半：工程师所做的工作是制造一种通用的图灵机，而正是通用计算机器的观念才是真正革命性的和更为本质的。按照戴维斯的说法，计算机实际上是逻辑机器，它的电路体现了几个世纪以来一大批逻辑学家所提出的观点之精华。当前，正当计算机技术以惊人的速度突飞猛进之时，正当我们羡慕工程师们令人瞩目的成就之时，我们很容易忘记那些逻辑学家，正是他们的思想使得这一切成为可能。从莱布尼茨到图灵，计算机的硬件和软件体现了几个世纪以来的一批逻辑学家所提出来的概念。本书讲述的就是这些位于计算机背后的思想层面的历史。它通过引人入胜的材料描写了这些天才的生活和工作，讲述了数学家们如何在成果付诸应用之前很久就已经提出了其背后的思想。文中语言生动而浅显，把这样一个相对枯燥的主题写得有声有色。通过阅读，读者可以在相当程度上了解

计算机是怎样工作的，其内部的算法是怎么一回事，从而消除对计算机的神秘感。同时还可以对西方文化的核心之一——逻辑或数学有更深的理解，并且造就一种敏锐的眼光和问题意识，认识到再复杂的东西其实也是由简单的东西根据一定的规则组合而成的。在普遍崇拜技术外表而忽视其深层本质的今天，这样的书就显得尤为难得和重要。希望您下次启动计算机时，脑海里浮现出来的不仅有工程师，而且还有那些伟大的数学家和逻辑学家们。

由于译者不是计算机方面的专家，翻译不当之处一定不少，恳请读者不吝赐教。中国科学技术大学的汪芳庭教授曾热情地回答过译者所提出的一些专业问题，陈宇也曾就此译稿提出过不少改进意见，这里谨向他们表示衷心的感谢！

张卜天

2004年8月于北京大学

[1] 伏尔泰《老实人》中的邦葛罗斯就是对这种莱布尼茨学说的一个讽刺。

[2] 1623年6月19日，布莱斯·帕斯卡出生于法国的克莱蒙费朗（Clermont-Ferrand）。他是概率的数学理论的奠基人之一，也是一位多产的数学家、物理学家和宗教哲学家。他于1643年前后设计并制造的计算装置使他享有盛誉。他于1662年去世。

[3] 尽管惠更斯的观点后来被普遍接受，但20世纪出现的量子物理学却表明，牛顿和惠更斯都是正确的，他们每个人都把握住了光的一个本质特征。

[4] π/4实际上是半径为1/2的圆的面积。

[5] 这一用莱布尼茨级数计算出的π/4的值是通过在一台486 33 MHz的PC机上编写和运行一个Pascal程序而得出的。把1000000项相加需要运行50秒的时间，把10000000项相加需要运行8分钟的时间。两年以后，这一程序又在一台奔腾200 MHz的机器上重新运行，所需时间分别被减少为4秒和40秒！

[6] 于是，确定体积和重心的问题属于第一类问题，而计算加速度和（在经济理论中）边际弹性的问题属于第二类问题。

[7] 表示积分的符号∫其实是字母“S”的变形，暗示“和”（sum）；类似的，符号“d”暗示“差异”（difference）。

[8] 阿里阿德涅（Ariadne）是古希腊神话中国王米诺斯（Minos）的女儿，曾给情人忒修斯（Theseus）一个线团，帮助他走出迷宫。——译注

[9] 这种印象来自库尔特·胡伯教授在狱中等待纳粹行刑期间写成的传记[参见参考书目中的（*Huber* ）一书]。他曾经支持过他在慕尼黑大学的学生们的努力，这些学生曾经组成地下团体“白玫瑰”，并且由于散发反纳粹传单而被杀害。现在的慕尼黑大学有一个以胡伯教授命名的广场（感谢班森·梅茨提供关于胡伯教授这一英雄角色的信息）。

[10] 伦敦一个类似机构的研究[Barret-Ducrocq]讲述了许多这样的悲惨故事。

[11] 布尔的方程xx=x可以与莱布尼茨的A⊕A=A相比。两者都是要把一个运算应用于两项，当把这种运算应用于一项和它自身之时，得到的结果是同一项。

[12] George Everest（1790—1866），英国人，曾任印度大地测量局总测量师。英国殖民者以他的名字命名珠穆朗玛峰。——译者

[13] 符号逻辑学会这一国际性组织为传播新的研究出版了两份季刊，并且定期举行会议。欧洲的逻辑学家有他们自己的年会。有关逻辑和计算机之间关系的新工作会提交给“计算机科学中的逻辑”和“计算机科学逻辑”国际年会。

[14] 中央党以罗马天主教会为旨归。它的“教皇绝对权力主义”指的是来自罗马的影响。

[15] 我曾应邀在1979年的一次纪念《概念文字》100周年的科学会议上做了一个报告，在那次报告中，我回顾了它对于计算机科学的重大意义。这是我初次研究计算机科学在逻辑学上的历史背景，也是以本书为最终成果的研究的开始。

[16] 在这两种缩写形式中，原文分别为英文单词和它的首字母。——译注

[17] 现在普遍认为，通过使用数值坐标，几何学也可以被还原为算术。然而，弗雷格一直认为几何学必须被单独考虑。感谢帕特丽夏·布兰切特强调了弗雷格思想的这个方面以及对这一节所提出的其他有益评论。

[18] 弗雷格在其定义“数”的失败尝试中也用到了同一思想。

[19] 注意，大括号{……}表示所列各项组成了一个集合。

[20] 也许是由于康托尔使用了希伯莱字母，有一则流传甚广的猜测说，康托尔是一个犹太人。事实上，他的父母都是基督徒，他是在路德教信仰中被抚养长大的。在纳粹统治德国期间，是否有犹太人祖先是决定其数学能否被接受的重要因素。的确，相信这一点是有某种理由的，他父亲曾经有一个犹太祖先在15世纪的最后10年中被驱逐出了葡萄牙。在1895年4月30日写的一封信中，康托尔对他为何使用希伯莱字母做出了解释：“这是因为在我看来，其他字母似乎[已经]用得太多了。”（感谢舍尔曼·施泰因给我看了这封信的一份复印件。）

[21] ω是希腊字母表中的最后一个字母，读作“欧米伽”。

[22] 1994年的诺贝尔经济学奖授予了三个人，其中两个是经济学家，另一个是数学家约翰·纳什。之所以把奖授予纳什，是因为他在1950年写的博士论文中发现有一个定理可以在经济学以及其他地方找到许多应用。在这篇论文中，纳什天才地运用了布劳威尔的不动点定理。

[23] 公元前44年，恺撒被一班反对君主制的罗马元老院议员刺杀，行刺者包括他最宠爱的助手、挚友和养子——马尔库斯·尤尼乌斯·布鲁图斯（Marcus Junius Brutus），当恺撒最终发现布鲁图斯也拿着匕首向他时，他绝望地说出了这句遗言，放弃了抵抗，身中23刀，倒在庞培的塑像脚下气绝身亡。这句话被广泛用于西方文学作品中关于背叛的概括描写。—译者注

[24] 1903年，20世纪最伟大的数学家之一约翰·冯·诺依曼出生于布达佩斯。他在一个富裕的家庭中长大，是一个少年天才，能够充分利用一切资源来培养自己的天分。他在纯粹数学和应用数学的广阔领域（包括数学物理和经济学）都有建树。1933年，他在普林斯顿高等研究院建立之初就成为它的一员，并且保持这个职位一直到1957年去世。在第二次世界大战期间，他参与研究了军事问题，其中包括洛斯阿拉莫斯的原子弹项目。这种一直持续到冷战时期的兴趣使他非常关注高级计算设备的发展。

[25] 当我于20世纪40年代末读研究生时，曾经极为幸运地聆听了这些大科学家的两次讲演。作

为科学阐释的例子，这两个讲演都算不上出色，但这当然不是最重要的。我们蜂拥到Fuld大厅（高等研究院总部所在地）去聆听这些传奇式的讲演。至于赫尔曼·外尔的情况，需要介绍日本数学家Kodaira所设立的一系列讲演。关于他的讲演，我记得最清楚的就是他在谈及数学思想时的愉悦之情。外尔的讲演组织得相当不好，尽管接下来的Kodaira讲演都属于清晰地进行数学阐释的范本。爱因斯坦的讲演来源于他为“统一场论”发现的一套方程，他所感兴趣的是由所谓的变分原理导出统一场论。他在黑板上写着字，仅仅当J.罗伯特·奥本海默（研究院院长）提醒他时间时才停下来，这时他已经完全背离了那个时代的主流。

[26]　卡尔纳普在耶拿大学获得了博士学位，在那里他师从弗雷格。卡尔纳普是那种被称为逻辑经验主义的哲学运动的领军人物。自1935年起，他一直在美国大学任教，首先是在芝加哥大学，然后是在加利福尼亚大学洛杉矶分校。汉斯·哈恩是哥德尔的论文指导老师。他在数学的不少领域里都做出了重要贡献，并对哲学问题有着浓厚的兴趣。

[27]　解释程序把一个程序的许多步骤逐步翻译成机器语言，并且在进行到下一步之前执行每一个步骤。编译程序则把整个程序翻译成机器语言。如此产生的机器语言程序可以作为一个独立条目运行，而不再需要编译程序。几乎所有的商业软件都是由编译程序生成的。

[28] 有材料说，她在一家名叫“夜蛾”（Der Nacht falter）的夜总会跳舞，“蛾”这个名字旨在暗示晚间幽暗的生物。另有人说她是一名芭蕾舞演员。

[29]　有理由相信，这对新婚夫妇曾经有过同去普林斯顿的计划。参见[Dawson]（在参考书目中），第128—129页。

[30]　这是应美国数学会之邀所做的具有很高声望的吉布斯年度演讲。我非常幸运地听了那次演讲，它对我关于数学基础的看法产生了深刻的影响。

[31] 1791年12月，查尔斯·巴贝奇出生于伦敦。他是一个颇有成就的数学家，他所在的群体试图把大陆数学思想传播到英国的大学中。他曾对机器运算产生了浓厚的兴趣，并且设想出一种机器——“差分机”，旨在有效地构造数学表。不久，巴贝奇突发灵感，构想出他那更为野心勃勃的分析机。1871年，未能完成这项计划的他满怀着痛苦和失望离开了人世。

[32]　正如大多数读者可能意识到的，英国的公立学校实际上是精英的私立机构。上一所公立学校在一个孩子通往成功的中产阶级职业的过程中是重要的一步。

[33]　法国、德国和美国的大学任命通常需要博士学位，而第二次世界大战前的英国学术界则很少要求这一点。

[34] 事实上，这时“计算机”一词的意思是：一个以计算为工作的人（通常是女性）。

[35]　图灵当然没有把它们称为图灵机，他当时把它们叫做a—机——“a”表示“自动”（automatic）。

[36] 注意，这一编码方案允许带子上的符号可以不是十进制数和□，而是像81118这样的符号。这就使得标记带子上的特定方格的符号可以出现，从而在返回时可以被找到。可以证明，使用这种附加的符号并不会增强图灵机的计算能力。还可以证明，使用十进制系统与图灵机所能做的事情没有关系。参见[Dav-Sig-Wey]（列在参考书目中），第113～168页。

[37] 对这一点不熟悉的读者，参见第4章“对角线方法”一节。

[38] 应当强调指出，如果真有一种把D的成员与非成员区分开来的算法，那么这些输入输出就不成其为问题了。毕竟，把输入的数交给一个人去执行那种形式的算法不会有问题，让她把输出的数以所希望的形式印在带子上也不会有什么问题。

[39] 例如，图灵说明了如何构造出能够产生0和1的字串以得到实数e和π的二进制表示的机器。对于标准数学中所出现的其他实数——整系数多项式方程的根甚至是贝塞尔函数的实零点，他也如法炮制。

[40] 阿隆佐·丘奇（1903—1995）在繁荣美国的逻辑学研究方面起到了关键性的作用。他创立了颇有影响力的《符号逻辑杂志》（Journal of Symbolic Logic），并且担任它的编辑长达40多年。另一位著名的美国逻辑学家斯蒂芬·克林（1909—1994）是丘奇的31位博士生之一（我也是其中之一）。

[41] 刊登丘奇对判定问题不可解性的证明的那一期《符号逻辑杂志》上还刊登了美国逻辑学家E.L.波斯特所写的一篇短文，这篇文章提出了一个相当接近于图灵的概念（[Davis 1]，第289～291页）。当我在纽约大学城市学院读本科时，波斯特是我的老师。

[42] 普林斯顿大学今天所在的建筑也被称为范氏大楼，从一英里以外的美国一号高速公路上看，它就像一座用水泥浇铸成的塔。

[43] 斯坦尼斯拉夫·乌拉姆（1909—1984）是重要的纯粹及应用数学家，他在许多数学分支上都有贡献，也是冯·诺依曼的好朋友。他有一种想法最终成就了一种推广集合论普通公理的重要方法，它有助于理解哥德尔关于连续统假设的工作。对于乌拉姆最重要的贡献——裂变聚变热核武器的基本设计，并不是每个人都赞同的。

[44] 可巧的是，这个车间就坐落在范氏大楼旁边的帕尔默物理实验室中。这两幢建筑之间还有一条过道。

[45] 数学家戈登·威尔希曼比图灵大6岁，他给图灵的设计加上了一个非常重要的部件，从而大大增强了它的能力。对破译“谜”的通信的技术细节感兴趣的读者可以参阅威尔希曼本人的叙述[Welchman]以及[Hodges]。[Hinsley]中收有解密工作的几位参与者对战争期间布莱奇利庄园中的生活的生动叙述。（括号中的书目参见书后所附的参考书目。）

[46] -*ismus* 是一个德文词尾，它的使用很像英文中的-ism。

[47] 今天计算机上的存储器是由被称为RAM的芯片制成的，它的意思是“随机存取存储器”。

[48]　我本人是在1951年春才知道计算机程序设计的，那时我开始为乌尔班纳—尚佩恩的伊利诺伊大学所研制的一台“约尼阿克”机ORDVAC编写代码。1954年夏，我写了一个使普林斯顿高等研究院的原始的“约尼阿克”机得以运转的程序（与莱布尼茨的梦毫不相干）。今天，那台计算机仍然保存在华盛顿的史密森学会。

[49] 1954年夏，我正在为冯·诺依曼的高等研究院计算机的基本数字运算指令集而绞尽脑汁，正在执行一种用于检验PA（定义见第6章）句子是否为真的算法，它包含加法，但不包括乘法。（计算逻辑领域的技术论文集的编辑们在前言中谈到我的项目时说：“1954年的一个计算机程序造就了计算机生成的第一个数学证明。”[Siek-Wright]p.ix.）我毫不怀疑，对我的目的来说，ACE指令集会合适得多。

第一推动丛书: 综合系列

The Polytechnique Series

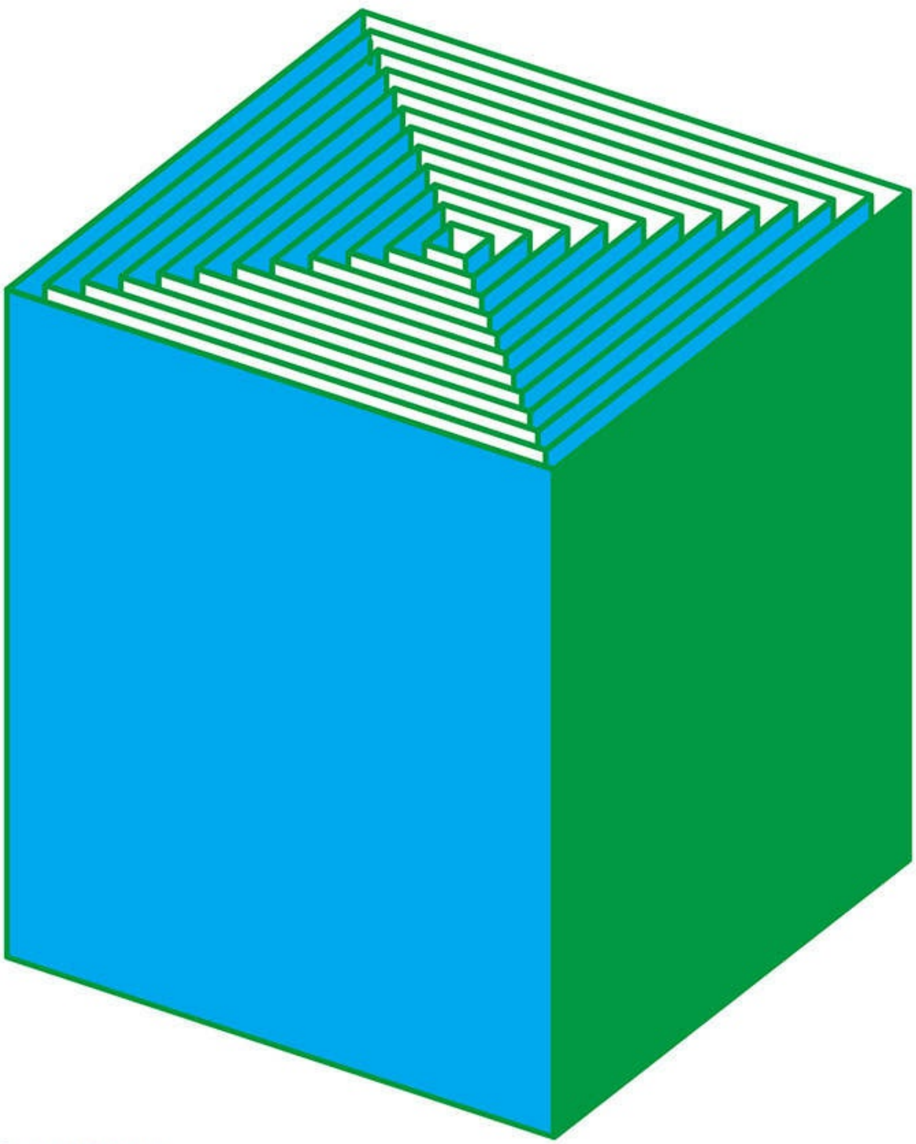
数学的意义

Meaning in Mathematics

[英] 约翰·查尔顿·波金霍尔 编
王文浩 译

John Charlton Polkinghorne

THE FIRST MOVER



湖南科学技术出版社

版权信息

数学的意义

著者：[英]约翰·查尔顿·波金霍尔

译者：王文浩

责任编辑：吴炜　孙桂均　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794352

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目 录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
各篇文章作者简介
引言
第1章 数学是一种发现还是一种发明？
评蒂莫西·高尔斯的“数学是一种发现还是一种发明？”
第2章 探索巴别数学图书馆
评马库斯·杜·索托伊的“探索巴别数学图书馆”
第3章 数学实在
评约翰·波金霍尔的“数学实在”
第4章 数学、大脑与物理世界
评罗杰·彭罗斯的“数学、大脑与物理世界”哥德尔定理与柏拉图主义
第5章 数学的理解
对彼得·利普顿的“数学的理解”的补遗
第6章 数学中的创造和发现
评玛丽·伦的“数学中的创造和发现”一文被感知的客观性
第7章 发现、发明和实在论：哥德尔和其他人关于概念实在性的观点
评迈克尔·德特勒夫森的“发现、发明和实在论”
第8章 数学与客观性
评斯图尔特·夏皮罗的“数学与客观性”
答复吉迪恩·罗森
第9章 数学对象的实在性
评吉迪恩·罗森的“数学对象的实在性”
第10章 我们从数学中得到的要比赋予它的多
评马克·施泰纳的“我们从数学中得到的要比赋予它的多”
参考文献

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

各篇文章作者简介

约翰·查尔顿·波金霍尔[1]（John Charlton Polkinghorne），英国高等爵士（KBE [2] ），英国皇家学会院士，剑桥大学皇后学院前院长，2002年度邓普顿奖（Templeton Prize）获得者，20多年来一直是科学与宗教对话方面的领袖级人物。1979年，他辞去了剑桥大学数学物理教授的职务，开始了后半生新的宗教生涯，并于1982年被任命为英格兰教会的牧师。作为英国皇家学会的院士，1997年他被女王伊丽莎白二世封为爵士。他在理论粒子物理学方面的著作甚丰，代表作有《量子力学：简明导论》（2002）。除此之外，他还主编或共同主编过4本书，并与迈克尔·韦尔克（Michael Welker）合著了《对永生的神的信仰：对话集》（2001）。在科学与神学的交叉领域，他先后写过19本书，其中包括《科学时代的上帝信仰》（1998年出版，是他在耶鲁大学特瑞讲席演讲的汇编），《科学与技术》（1998），《信仰、科学与理解》（2000），《真理的传递：技术与科学之间的交流》（2001），《希望之神与世界末日》（2002），《与希望同在》（2003），《科学与三位一体：基督教如何面对实在》（2004），《探索实在：科学与宗教的共赢》（2005），《量子力学与技术：出人意表的亲缘关系》（2007），《从物理学家到牧师》（2007），《科学语境下的技术》（2008），以及与尼古拉斯·比尔合著的《真理之问：有关上帝、科学和信仰等问题的51个回答》（2008）等。

迈克尔·德特勒夫森（Michael Detlefsen），巴黎圣母院大学麦克马洪—汉克哲学讲席教授，巴黎第七大学—狄德罗大学和南锡第二大学特聘客座教授。自2007年以来，一直担任法国国家研究机构（ANR [3] ）的资深主席。主要学术兴趣在数学和逻辑学的历史研究和哲学研究。目前正在进行的工作包括与蒂莫西·麦卡锡（Timothy McCarthy）合写一本有关哥德尔不完备定理的书，另外几项工作涉及数学证明的理想。

马库斯·杜·索托伊（Marcus du Sautoy），牛津大学数学教授和“科学的公众理解”西蒙尼讲席教授，新学院院士。他的学术领域主要涉及群论和数论。他在普及数学知识方面非常著名。获得过2001年度伦敦数学会颁发的贝维克奖（Berwick Prize）和2009年度英国皇家学会的法拉

第奖（Faraday Prize）。他是英国广播公司（BBC）多部电视专题片和广播系列节目的主讲者，先后出版了3本科普读物：《素数的音乐》（2003）[4]，《发现空想：数学家的对称之旅》（2007）和《数字之谜：日常生活中数学奥德赛》（2010）[5]。

蒂莫西·高尔斯（Timothy Gowers），英国皇家学会院士，剑桥大学罗斯·鲍尔（Rouse Ball）数学讲席教授，剑桥三一学院院士。1998年因在泛函分析与组合学领域的突出成就荣获菲尔兹奖。在此之前，他还荣获过伦敦数学学会颁发的小怀特海奖（Junior Whitehead Prize）和欧洲数学学会奖。他出版的书有：《数学：简明导论》（2002）和《普林斯顿数学手册》（2008）（主编）。2009年，他推出了“博学项目”（Polymath Project），采用博客的评论功能来进行数学协作研究。

玛丽·伦（Mary Leng），英国利物浦大学哲学讲师。她的研究重点是数学哲学，特别是数学在实证科学中应用产生的问题。玛丽·伦博士现为美国加州大学欧文分校的逻辑与科学哲学客座研究员，此前她先后获得过多伦多大学人文学科的博士后奖学金、剑桥大学圣约翰学院的研究奖学金（为期4年），以及担任加拿大英属哥伦比亚大学彼得·沃尔（Peter Wall）高等研究院的访问学者。2007年，她与亚历山大·帕赛和迈克尔·波特（Alexander Paseau and Michael Potter）合编了《数学知识》，2010年，她出版了《数学与实在》一书。这两本书均由牛津大学出版社出版。

彼得·利普顿（Peter Lipton，1954—2007），剑桥大学汉斯·劳辛（Hans Rausing）讲席教授和科学史与科学哲学系主任。他是剑桥大学国王学院的研究员。尽管他的大部分研究都是关于科学的论述和推理，但他的兴趣广泛延伸到诸多哲学领域。他曾以医学科学院研究员的身份担任了《科学史和科学哲学研究》杂志的顾问编辑，主编了《理论、证据与解释》（1995），主编和合作主编了《科学史和科学哲学研究》的3期特刊。他是《最佳解释推理》一书的作者（1991年第一版，2004年再版）。

罗杰·彭罗斯（Roger Penrose），功绩勋章获得者（OM[6]），英国皇家学会院士，牛津大学罗斯·鲍尔（Rouse Ball）数学讲席终生荣誉教授，牛津大学沃德姆（Wadham）学院荣誉研究员。他在数学物理，特别是在广义相对论、量子理论基础和宇宙学等方面的原创性、宽厚的基

础性工作而广受赞誉。他还著书探讨基础物理学与人类意识之间的联系。作为皇家学会院士，美国国家科学院外籍院士和欧洲科学院院士，彭罗斯教授于1994年被女王伊丽莎白二世授以爵位以表彰他在科学上的贡献，并于2000年荣获英国功绩勋章。他先后写了（包括与人合著）10本书，包括《皇帝新脑》（1989）[7]，本书获1990年度科学图书奖；《心灵之影：探索鲜为人知的意识科学》（1994）；《大的，小的和人的心灵》（1997）。除了论述意识问题的书之外，他还为更广泛的普通读者写过其他一些书，包括与斯蒂芬·霍金合著的《时空本性》（1996）[8]；《通向实在之路：宇宙法则完全指南》（2004）[9]和《时间的轮回：一种异乎寻常的新宇宙观》（2010）等。

吉迪恩·罗森（Gideon A.Rosen），普林斯顿大学斯图亚特哲学讲席教授，人文学院院长。研究方向为形而上学、认识论、数学哲学和道德哲学。他是新西兰奥克兰大学的客座教授，并拿到梅隆基金会设在纽约大学法学院的新方向研究基金，出任全球行政法项目的Hauser研究员。罗森教授（与约翰·P.伯吉斯）合著了《没对象的学科：唯名论数学的解释战略》（1997）一书，该书由牛津大学出版社出版。

斯图尔特·夏皮罗（Stewart D.Shapiro），美国俄亥俄州立大学奥当奈（O 'Donnell）哲学讲席教授，圣安德鲁斯大学教授级研究员。他的研究和写作主要集中在数学哲学、逻辑、逻辑哲学和语言哲学等方面。他先后取得过多种人文研究资助和美国学术协会的资助，他还荣获过俄亥俄州立大学的学术成就奖和俄亥俄州立大学杰出学者奖。夏皮罗教授曾是《符号逻辑杂志》的编辑，主编过5期特刊和3本著作，其中包括《牛津逻辑和数学哲学手册》（2005）。他还在牛津大学出版社出版了4本著作：《无基础主义的基础：以二阶逻辑为例》（1991年初版，2000年再版）；《数学哲学：结构和本体论》（1997年初版，2000年再版）；《对数学的思考：数学哲学》（2000）[10]和《含糊的语境》（2006）。他还为牛津大学出版社编写了《哲学家的逻辑》（暂定名）新教科书。

马克·施泰纳（Mark Steiner），耶路撒冷希伯来大学哲学系教授。他的研究领域是科学哲学，但更关注数学哲学。他的研究包括对维特根斯坦（Ludwig Wittgenstein）的数学哲学的评述。出版的著作有：《数学知识》（1975）和《作为一个哲学问题的数学适用性》（1998）。他将鲁文·阿古什维兹（Reuven Agushewitz，一位立陶宛出生的犹太法典

《塔木德经》学者）用意第绪语写的*Emune un Apikorses* （1948年版，这本书攻击了历史上各种形式的哲学唯物主义）翻译成英文，取名为《信仰和异端邪说》（2006）。现在他正在将休谟的《人性论》由英文译成希伯来文。

引言

约翰·波金霍尔

数学到底是一种由行家施展身手来表演如何化解难题的高度复杂的智力游戏，还是数学家在探索数学实在这一独立领域过程中所带来的发现？为什么这个看似抽象的学科能够提供打开物理宇宙深层秘密的钥匙？如何回答这些问题将明显影响着我们对实在的形而上的思考。在冈道尔夫堡和剑桥召开的两次跨学科专题讨论会上，数学家、物理学家和哲学家们对这些问题进行了探讨。本书以周详的形式再现了每位与会者在会议热烈讨论中所展现的风采。文章尽力保持这样一种平衡：既反映进行这种讨论所需的思想精确性，又照顾到准备在此领域做出一番事业的非专业读者的可读性。

剑桥大学科学哲学教授彼得·利普顿参加了第一次会议，并有精彩发言。但不幸的是，这之后他溘然长逝，对此我们感到非常难过。所有与会者有一个共同心愿：将本书作为我们对这位尊敬的学者和谦和、富于启迪的同事的美好追忆。

本书的前两章由数学家蒂莫西·高尔斯和马库斯·杜·索托伊撰写。他们能够充分利用长期从事数学研究的丰富经验来阐述问题。高尔斯特别重视“发明”和“发现”这两个词是如何被数学界实际运用的。他的结论是，当导致重要结论的论证基本上只有唯一一条途径时，用“发现”来说明似乎是恰当的。而如果存在多条清晰的论证途径时，则人们更愿意用“发明”一词来形容。杜·索托伊描述了在洞察一个事件时灵感闪现的情形，这是这样一种经验：可以确信，有待识别的东西早就“已经在”那儿等待被发现了。

接下来的两章由数学物理学家约翰·波金霍尔和罗杰·彭罗斯撰写。波金霍尔旨在通过对哥德尔不完全性和人类数学能力进化的论述来捍卫数学实在。两位物理学家都非常看重数学在他们做出发现过程中所扮演的角色。彭罗斯认为，哥德尔不完全性意味着有意识的思想要远比神经

网络计算来得复杂。

其余章节由哲学家执笔。彼得·利普顿撰写的一章以短文呈现，以彰显他对第一次研讨会做出的贡献。这篇文章讨论了知识、理解和解释等概念，强调了他认为这些概念在科学和数学之间应用的差异。斯图尔特·夏皮罗帮忙为本文提供了一个附录，说明本次讨论的一些方法可能会得到进一步扩展。玛丽·伦对那种发现的感觉——许多数学家论证认为必然由此导致柏拉图的数学实在的观点——持否定态度。相反，她认为，这种感觉可以理解为出自逻辑上的必然性。迈克尔·德特勒夫森则对古代和现代围绕发明或发现的争论进行了广泛的调查。他对哥德尔著名的数学“知觉”与感性知觉之间的类比给予了谨慎的批评。斯图尔特·夏皮罗认为，数学是一种人类活动，其传统源自人类的选择。在他看来，关键概念是“认知律令”。这个概念用来说明不同的人做同样的计算所取得的结果应有必然的一致性这一现象。他认为这一观点鼓励人们从发现的角度去看问题。吉迪恩·罗森探讨了这样一种观念：数学的地位相当于他所谓的“有条件的实在论”。他将这一判断描述成对数学作为“形而上学上第二等”的一种裁决，因为它依赖于更基本的逻辑事实。最后，马克·施泰纳将我们领向笛卡儿而不是柏拉图。他强调，数学似乎能够提供某种“剩余价值”，允许数学家超越公理（数学家自己则将这种超越称为“深入”的品质）。

本项研讨会的一个特点是讨论氛围的活泛和透彻。与会者希望本书能将这种气质传递给读者，因此我们对每篇文章都附上一篇由其他与会者撰写的短评。我们相信，这些评论是正文报告的一个重要组成部分，它反映了研讨会带来的启发性和挑战性。

研讨会的两次会议均得到了约翰·邓普顿基金会（John Templeton Foundation）的支持。所有与会者对这一慷慨资助表示由衷的感激。我们特别要感谢基金会的玛丽·安·迈尔斯博士，她在组织协调方面提供了大力帮助，并对会议议题表现出浓厚兴趣。

第1章　数学是一种发现还是一种发明？

蒂莫西·高尔斯

本章标题是一个著名的问题。事实上，也许这个问题有点过于出名了：不断有人提出这个问题，但怎么作答都不能令人满意。在形成本书的讨论中，大家推举我来回答这个问题。由于大多数参与讨论的都不是研究数学的专家，因此希望我能从数学家的角度来处理这个问题。

提出这个问题的一个原因似乎是人们希望用它来支持自己的哲学观点。如果数学是一种发现的话，那便意味着原本就有某种东西在那里等待数学家去发现，这种认识似乎支持了柏拉图主义的数学观点；而如果数学是一种发明的话，那么它则为非实在论关于数学对象和数学真理的观点提供了某种论据。

但在得出这样一个结论之前，我们需要从细节上充实论据。首先，当我们说数学的某项内容被发现时，我们必须十分清楚这指的是什么，然后我们必须在这个意义上解释清楚为什么能够得出这一结论（这套程式被称为柏拉图式论证）。我自己并不认为这套做法能够贯彻到底，但它至少从一开始就试图阐明这样一个不争的事实：几乎所有数学家在成功证明某个定理时都会感到好像他们有某种发现。我们可以用非哲学的方式来看待这个问题，这里我正是尝试这么做的。例如，我会考虑是否存在某种可识别的东西，以便鉴别哪些东西看上去像是数学发现，哪些更像是数学发明。这个问题部分属于心理学范畴的问题，部分属于是否存在数学陈述的客观性的问题，即属于解释某个数学陈述是如何被感知的问题。要想论证柏拉图的观点成立，我们只需要指明存在某些被发现的数学事实就足矣：如果事实证明，存在两大类数学，那么我们或许就能够理解这种区别，对何为数学发现（而不是单纯的数学结果）做出更精确的定义。

从词源上说，所谓“发现”通常是指当我们找到了某个早已在那儿但我们此前不知道的东西。例如，哥伦布对美洲的发现（尽管人们出于其

他原因对此大可质疑），霍华德·卡特于1922年发现了图坦卡蒙的墓，等等。尽管所有这些发现并非我们直接观察到的，但我们依然能够这样说。例如我们都知道是J.J.汤姆孙发现了电子。与数学关联更强的是如下事实的发现——例如我们可以确切地说，是伯恩斯坦和伍德沃德发现（或对这一发现有贡献）了尼克松与水门入室盗窃案有关。

在所有这些情形里，我们都观察到一些引起我们注意的现象或事实。因此有人可能会问，我们是否可以将“发现”定义为从未知到已知的转变过程。但有不少事例表明，事实并非如此。举例来说，喜欢做填字游戏的人都知道这样一个有趣的事实，单词“carthorse（大马）”和“orchestra（乐队）”属于一对字母换位词。我相信肯定是某个地方的某个人最先注意到这个事实，但我宁愿将它称为“观察”（我用“注意到”这个词来描述这一事实）而不是“发现”。为什么呢？这是因为“carthorse”和“orchestra”这两个词我们每天都用，它们之间是一种简单关系。但是为什么熟悉的单词间关系我们不能称之为发现呢？另一种可能的解释是，一旦这种关系被指明，我们很容易验证它的成立，我们没必要从美国跑到埃及去宣讲这一事实，也没必要通过做精密的科学实验予以验证，或设法获取某个秘密文件才能知晓。

至于谈到柏拉图式论证的证据，“发现”和“观察”的区别不是特别重要。如果你注意到某个事实，那么这个事实一定在你注意到它之前已经在那里了，同样，如果你发现了某个事实，那一定是在你发现之前它就存在了。因此我认为，观察事实属于某种发现而不是一种根本不同的现象。

那什么是发明呢？我们做的什么样的事情属于发明呢？机器是一个很好的例子：谈到蒸汽机，或飞机，或移动电话，我们说这些是发明。我们还认为游戏属于发明，例如英国人发明了板球。我更想指出的是，“发明”是以适当的方式来描述所发生的事。艺术为我们提供了这方面的一些更有趣的例子。人们从来不会说某个艺术品是被发明出来的，但会说发明了某种艺术风格或技巧。例如，毕加索不是发明了《阿维尼翁的少女》（*Les Desmoiselles d’Avignon* ），但确实是他和布拉克发明了立体派绘画艺术。

从这些例子我们得出一种共识，我们发明的东西往往不是单个对象，而是生产某类对象的一般方法。当我们说到蒸汽机的发明时，我们

不是在谈论某台特定的蒸汽机，而是一种概念——一种将蒸汽、活塞等东西巧妙地结合起来用以驱动机器的设计，它能够导致许多蒸汽机的建造。同样，板球是一套规则，它可以带来各种形式的板球运动，立体派则是一种对各种立体派绘画的一般性指称。

有人将数学发现这一事实看作柏拉图学派数学观的证据，其实他们试图表明的是，某些抽象实体具有独立存在的属性。我们认可体现这些抽象实体真实性的某些事实，与我们接受具体事物真实性的某些事实有大致相同的原因。例如，我们认为存在无穷多个素数这一陈述就是一种真实的事实，这是因为的确存在无限多的自然数，并可确信，这些自然数中确实存在无限多个素数。

有人也许认为，抽象概念是一种独立存在这一点也可以作为“数学是一种发明”这一观念的证据。确实，我们有关发明的很多例子都以某种重要方式与抽象概念相关联。前述“蒸汽机”便是这样的一个抽象概念，板球规则也是如此。绘画中的立体派是一个比较麻烦的例子，因为它没有那么精确的定义，但它无疑是具体的而不是抽象的。但我们发明这些概念时为什么不说这些抽象概念是一种存在呢？

一个原因是，我们认为独立存在的抽象概念应该是永恒的。因此，在英国人发明板球规则时，尽管这些规则属于抽象领域并成为一种存在，但我们不倾向于认为它们是永恒的。更诱人的一种观点是，他们是从巨大的“规则空间”里选择了板球规则，这个规则空间包含了所有可能的规则集（其中大部分规则会引起可怕的游戏）。这种观点的缺陷是，它用大量垃圾概念充斥了抽象领域，但实际情形也许真的是这样。例如，数空间显然包含所有的实数，但其中除了“可数的”这个子集外，其他的都无法定义。

反对“我们发明抽象概念从而使它变成存在”的另一种论证认为，我们发明的概念不是基本的，它们往往是对其他一些（抽象或具体的）更简单的对象的处理方法。例如，板球规则的描述就涉及包含22名球员、1个球和2个球门的一组概念之间的约束。从本体论的角度看，球员、球和球门显然比约束它们之间相互关系的规定更基本。

前面我提到过，谈到某一件艺术品时我们通常不会用“发明”一词来指称。当然我们也不会说“发现”了它，而是通常用“创造”这个词来指

称。大多数人在被问到这个问题时都会认为，“创造”一词在这里其词意接近于“发明”而不是“发现”，正像“观察”一词的词意接近于“发现”而非“发明”。

这是为什么呢？是这样的：在这两种情况下，变得存在的那件东西原本有许多任意性。如果我们可以将时钟拨回到板球被发明出来之前，让世界重新演化一遍，我们很可能会看到发明出一种类似的游戏，但其游戏规则不太可能与板球比赛规则完全相同（有人可能会反驳说，如果物理定律是确定的，那么这个世界应当精确地按照它第一次演化时那样演化。在这种情况下，它重新演化时只会做一些小的随机变化）。同样，如果有人在毕加索刚创作完《阿维尼翁的少女》后便不小心毁坏了它，迫使毕加索不得不重新开始创作，那么他创作的可能是一幅类似但不完全相同的画。与此相反，如果没有哥伦布，那么也会有其他人发现美洲，而不是在大西洋的另一边发现的只是一块巨大的、面积大致相同的陆地。而单词“carthorse”和“orchestra”的有趣性与谁是第一个观察到这一点无关。

有了这些思想上的准备，现在我们回到数学上来。同样，我们先看一些人们常列举的著名例子将有助于我们对问题的理解。我先列举一些发现、观察和发明的事例（我无意设定这样一种场景，好像我可以确定地表示某些数学问题是创造出来的），然后尝试着解释为什么每个事例会是按这种方式来描述。

数学上几个知名的发现方面的例子是：二次方程有通解公式但五次方程则没有类似的公式；存在大魔群；存在无穷多个素数。稍微观察一下即可知，小于100的素数的个数是25；3的各次幂的最后一位数字构成序列3，9，7，1，3，9，7，1，…；数字10 001可分解为73乘以137。水平稍高一点儿的事例则有：如果你通过设$z_0=0$，$z_n=z^2_{n-1}+C$（对每个$n>0$）定义了一个无穷的复数序列z_0，z_1，z_2，…则所有复数C的集合（假定序列不趋于无穷大，这个集称为曼德布罗特集）具有显著的复杂结构（我将这个例子归于中等数学水平是因为，虽然曼德布罗特和其他人几乎是由于偶然而无意中发现了它，但它已经在动力系统理论中具有根本的重要性）。

另一方面，人们常说牛顿和莱布尼茨各自独立地发明了微积分（用这个例子的过程还有点巧合。我原本考虑过这个例子。那天在我写这一

段时，恰好电台在播送关于他们的优先权纠纷的节目，用的词就是“发明”）。人们有时也会谈论某些数学理论（而不是定理）的发明。说格罗滕迪克（A.Grothendieck [11] ）发明了概型理论（theory of schemes），这听起来一点都不荒谬，虽然人们也可能同样会用“引入”或“发展”来描述这一理论。同样，这三个词也都被用来描述P.J.科恩（P.J.Cohen）的力迫法（forcing method），他用这一方法来证明连续统假说的独立性。 [12] 这里我们感兴趣的是，“发明”“引入”和“发展”这三个词都暗示了这样一点：某些一般性方法应运而生。

有可能存在争议的一个数学对象是“i”或更一般的复数系统。复数是一种发现还是一种发明？或者说，数学家通常提到复数进入数学领域时用的是“发现”类型的词还是用“发明”类型的词？如果你在Google上输入词组“复数的发明”和“复数的发现”，你会得到大致相同的点击次数（二者均在4500～5000），所以这个问题似乎没有明确答案。但这也是一个有用的数据。类似的一个例子是非欧几何，虽然“非欧几里得几何的发现”与“非欧几何的发明”的点击次数的比例约为3:1。

另一种不明确的情形是证明：证明是一种发现还是一种发明？有时候证明似乎很自然——数学家常讲，一个陈述的“（逻辑上）正确的证明（right proof）”并不意味着它是唯一正确的证明（correct proof），而只是表明它是一个能真正解释为什么这一陈述是对的证明——“发现”这个词显而易见可用于这一情形。但有时候人们谈到类似的东西时却感到下面这样的表述更恰当，譬如说，“猜想2.5在1990年首度获得证明，但在2002年，史密斯给出了一个巧妙且非常简短的证明，这一证明实际上建立了一种更一般的结果。”在这句话里，人们可以用“发现”来替代“给出（came up with）”一词，但后者更好地刻画了这样一层意思：史密斯的方法只是众多有可能提出的方法中的一种，但史密斯并不是简单地纯属偶然地找到了这种方法。

让我们来小结一下上述观点，看看数学中的哪些部分可以归结为发现、发明或不能明确用二者来描述，并能否给予解释。

非数学的例子表明，当发现者对观察对象或事实不能控制时，我们通常用“发现”和“观察”来描述这一发现过程。而当对象或程式具有许多可由发明者或设计者选择的特性时，那么我们就用“发明”或“创造”来描述这一对象或程式的诞生过程。由此我们也得出了对这两类过程的一些

更精细但不那么重要的区别：“发现”往往比“观察”更重要，但较不易事后验证。而“发明”往往比“创造”更具一般性。

当我们谈论数学时，这些区别还会继续保持上述大致相同的形式吗？前面我提到二次方程解的通项公式被发现的例子。当我尝试说“二次方程解的通项公式的发明”这句话时，我发觉我不喜欢这么描述，确切的原因是，ax^2+bx+c的解是数$\frac{-b\pm\sqrt{b^2-4ac}}{2a}$，无论是谁最先推导出这个公式，但这一公式的最终形式是什么样子是没有任何选择余地的。当然记述公式的符号可能会有不同，但那是另一回事。我不想在此讨论两个公式“本质上相同”是什么意思，这里我只需简单地说，公式本身是一个发现，但不同的人会用不同的方式来表达它。然而当我们来考察其他的例子时，这种担心会再次出现。

五次方程无通解就是另一个简单例子。它是指“五次以及高于五次的常系数代数方程没有由方程系数经有限次四则运算和开方运算确定的求根方法”[13]，对此阿贝尔也无法改变。因此，他的著名定理就是一种发现。然而，他的证明方法的具体细节可以被看作发明，因为后来有了很不相同的证明方法。这其中特别值得指出的是伽罗瓦的与此密切相关的工作，他发明了群论（“群论的发明”这句话在Google上有40 300条，相比之下，“群论的发现”只有10条）。

大魔群则是一个更有趣的例子。在费希尔（B.Fischer）和格里斯（R.L.Griess，Jr.）于1973年预言存在这种群后，它第一次进入了数学领域。但这句话意味着什么呢？如果他们根本不能给出具体的大魔群，难道就意味着它不存在吗？答案很简单：他们预言，存在一种具有特定显著特性的群（其中的一个显著特性是它元素巨多，大魔群由此得名[14]），而且这种群是唯一的。因此，说“我相信存在大魔群”只不过是“我相信存在一种具有这些惊人特性的群”这一语句的简短截说，名曰“大魔群”实际上指的是一个假设性的实体。

大魔群的存在性和唯一性的确切证明直到1982年和1990年才分别完成，在此之前我们不是很清楚是否应该把这个数学上的进展称为发现或是发明。如果我们略去其中的故事，将这17年浓缩为一瞬间，那么我们满可以这么说，大魔群一直就在那里，等待着群论数学家来发现。或许有人甚至可以添加一个小细节：早在1973年，人们就开始有理由假设它的存在，并且最终在1982年终于碰上了它。

那么这种“碰上”是怎么发生的呢？格里斯并没有用某种间接的方式证明大魔群必定存在（虽然这样的证明在数学上是可能的），而是构建了这种群。这里我像所有的数学家一样用“构建”这个词。为了构建大魔群，格里斯构建了一个辅助对象——一种现被称为格里斯代数的复杂的代数结构，并且证明了这种代数的对称性构成一种具有所需特性的群。然而，这不是获取大魔群的唯一方式，还存在能够产生具有相同属性的群的其他结构，因此从结果的唯一性来看，它们是同构的。如此看来，格里斯在建立大魔群的过程中有一定的控制权，甚至他最终想做到哪一步也是事先确定的。有趣的是，在Google上“大魔群的构造”这句话比那句“大魔群的发现”更流行（8290:9），但如果你把前一句改为“大魔群的这种构造”，那就变得不流行了（仅6条），这反映了一个事实，即大魔群有许多不同的构造。

有人可能会问的另一个问题是这样的。如果我们谈论大魔群的发现，我们是在谈论一个对象（大魔群）的发现，还是一个事实（即存在一种独特的、具有特定属性的群这一事实）的发现？当然，后一个设问对群理论家实际所做的工作是一个更好的描述，“构造”这个词在描述他们如何证明这一陈述的存在性时要比“发现”更准确。

先前列举的其他一些发现和观察事例表现得更直接，因此不赘述了。我们来看有关发明的例子。

在数学上，直接用“发明”这个词的情况多是指一般理论和技术的产生。这当然也包括微积分，但微积分不是一个对象或一个简单的事实，而是大量事实和方法的集成，如果你熟悉微积分，它将极大地提高你的数学能力。这种情形也包括科恩的力迫法技术。同样，这里也牵涉诸多定理，但我们真正感兴趣的是这一方法在证明集合论的独立表述中所表现出的一般性和普适性。

前面我提到过发明者在他们的发明过程中有一定的控制权。这也适用于这些例子：哪一种数学表述是微积分的一部分，这并没有明确的标准，同样，我们也可以有许多方法来表述力迫法原理（前面我已提到过对科恩“力迫法”原创思想的许多广义化、修改和扩展）。

复数系统又是怎样的一种情形呢？乍一看，这一点都不像一个发明。然而毕竟它被证明是唯一的（满足从a+bi到a-bi的同构），它是一

个对象而不是一种理论或技术。那么为什么人们不时称它为发明，或至少是觉得称它为发现会显得不那么自然呢？

对这个问题我没有成熟的答案。我认为其中的原因在于它带来的困难有点像大魔群——人们可以“构建”不止一种复数方式。构建复数的方法之一是运用类似于历史上复数最初被构建出来的方式（我的数学史知识不是很深厚，所以我不说二者的相似性到底有多接近）。人们简单引入一个新符号，i，并宣布它的性态就像一个真正的数，不仅遵从所有通常的代数规则，并具有额外的属性，$i^2=-1$。由此可以推断出

$$(a+bi)(c+di)=ac+bci+adi+bdi^2=(ad-bd)+(ad+bc)i$$

和许多可用于建立复数理论的其他事实。第二种构建方法（以后再做介绍，它表明复数系与实数系是一致的）是用一对有序的实数（a, b）来定义复数，并规定这些有序对满足以下给出的加法和乘法规则：

$$(a, b)+(c, d)=(a+c, b+d)$$

$$(a, b)(c, d)=(ac-bd, ad+bc)$$

在建立严格数系的大学课程里，经常使用的是第二种方法。业已证明，这些有序对构成了给定的两种运算下的一个域，最后我们可以说，“从现在开始，我将写成a+bi而不是（a, b）”。

我们对复数怀有这种两可心理的另一个原因是，它们不像实数那样给人真实的感觉（这一点从这些数的命名即可细微地反映出来）。我们可以直接将实数与时间、质量、长度、温度等的数量词联系起来（虽然在此我们无须用到实数系的无限精度），就好像它们具有一种被我们观察到的独立的存在属性。但是，我们不能将这种方式运用到复数上。相反，我们进行复数运算时总感觉到有点像在做游戏——如果-1存在平方根，想象一下会发生什么。

但是为什么在这种情况下我们还是会说发明了复数呢？原因是这个游戏比我们期望的更有趣，它在数学上已经对实数甚至整数产生了巨大影响。虽然发明i只是一个小小的举动，但游戏已失去控制，我们无法预料后果。[这种情况的另一个例子是著名的康威（Conway）生命游戏。康威设计了一个遵从一些简单规则的游戏，这款游戏与其说是发现

不如说是发明更确切。游戏一经开始，他便发现他实际上创造了一个充满意想不到现象的世界。其中的大多数现象在其他人看来明显可以说是发现]。

为什么“非欧几里得几何的发现”要比“非欧几里得几何的发明”更易被接受？这个问题之所以有趣，是因为我们有两种方式来看待这个主题：一个是公理性质的；另一个则具体些。将非欧几何看作重大发现的人们看到的是，不同的公理体系（在这里是平行线假设被替换为这样一种陈述——允许一条直线有几条通过任一给定点的平行线）是相容的。另一方面，我们又可以把非欧几何看作一种模型的构造。公理系统在这种模型里均成立。严格来说，要谈前一种方式我们就需要后一种方式，但如果我们从细节上探索这些公理的后果，并证明了所有有趣的定理均无矛盾，那么这些事实就会成为这种相容性的一种令人印象深刻的证据。可能是因为我们对理论相容性的兴趣大于对具体模型选择的兴趣，再加上这样一个事实——任意两个双曲平面模型等距，所以我们通常才将非欧几何称为发现。然而，欧氏几何让人（错觉地）感到要比双曲几何更“真实”，而且没有哪一种双曲几何模型看上去最自然。这两个事实可以用来解释为什么“发明”这个词有时会被用来形容非欧几何。

我的最后一个例子是关于证明的。前面我说过，证明根据其性质既可以称为发现也可以称为发明。当然，能用来指称的绝不是仅有这两个词或短语，人们可能还会用“想到”“找到”或“闪现出”等词语来形容。人们往往将证明视为对象而不是过程，侧重于要证明的东西，譬如这样的句子：“经过很长的推导，我们最终证明/建立/显示/表明了……”证明的这种特点再一次表明，当作者没有选择余地时，我们用“发现”来指称一个证明过程；如果可以有多种选择的话，就用“发明”来指称。人们可能会问，选择从何而来？这本身是一个有趣的问题，这里我仅指出这种选择和任意性的一个来源。证明通常需要我们给出存在特定的数学对象或结构（无论是主要陈述还是一些中间引理），而通常情形下问题中的这些对象或结构远非独一无二的。

在从这些例子得出结论之前，我想简要地讨论一下这个问题的另一个方面。以上我主要是从语言的角度来谈论问题，但正如我在开始时提到的那样，人们在选择用词时还有很强的心理因素在起作用：当一个人在做数学研究时，研究工作有时给人的感觉更像是发现，有时则更像是发明。这两种经验之间的区别是什么？

因为我比其他人更熟悉自己的研究经历，故我讲述一下我自己的经验。在20世纪90年代中期，我开始了一项研究。其实我很早就以这种或那种方式在琢磨这个问题。我在想，我应该能给一个定理一种比已有的两种证明方法更简单的证明。最后我找到了一种证明方法（这里我很自然地使用了这个词），尽管它不是更简单，但它给出了新的重要信息。找到这个证明的过程让人感觉到更像是发现而不是发明，因为当我接近完成这一证明时，论证结构中已包含了许多在证明开始时我甚至没有预料到的要素。此外更为明显的是，有一大堆密切相关的事实可用于一种自洽但尚未被发现的理论。（在这个阶段，它们还不是已被证明的事实，而且不总是准确陈述的事实，唯一明确的就是“有东西”值得研究。）我和其他几个人努力发展了这个理论，终于使得定理已被证明它甚至不像15年前作为猜想那样得到表述。

为什么这种感觉如同发现而不是发明？我们可以再次从它是否与可控制相联系这一点来理解。当时我没有从一大堆可能性中选择事实的余地，相反，某些陈述是以明显自然的和重要的方式呈现出来的。由于理论还在进一步发展中，因此哪些事实具有核心意义哪些属于边缘性质还不太清楚。而从这个角度看，研究过程又像是一种发明过程。

几年前我有了不同的经验。我找到了一个有关巴拿赫空间理论中一个古老猜想的反例。为此我构建了一个复杂的巴拿赫空间。在此过程中，我的感觉有时候像是在搞发明——我有任意选择的余地，许多其他反例随后被发现；有时候又如同发现——我做的很多工作都是回应问题本身提出的要求，而且感到这样做是很自然的事情。其他人也独立发现了一个非常类似的反例（甚至后续的例子中使用了类似的技术）。所以，这又是一种有待分析的复杂情形，说它复杂原因很简单，我有多大的控制权就是一个复杂的问题。

从所有这些例子中，以及从我们平时貌似自然地对待它们的情形里，我们应该得出什么样的结论呢？首先，很显然，我们最初提出的问题是相当人为的。就是说，开始时认为的所有的数学要不就是发现要不就是发明的想法是荒谬的。但是，即使我们考察了数学某个具体部分的起源，我们还是不会硬性使用“发现”或“发明”这样的语词，我们并不经常这么用。

当然，事情都有各种可能性。有些数学研究给人感觉像是发现，而

另一些则更像是发明。到底哪些属于前者哪些属于后者这并不总是容易说清楚，但似乎有一点可用于鉴别：这就是研究过程中的可选择性。这一点，正如我前面所说的，甚至有助于解释为什么有些可疑的情形是值得怀疑的。

如果这是正确的（也许还有待细化），那么我们可以从中得出什么样的哲学结论呢？在开始时我提过，像“数存在吗？”或是“数学陈述之所以正确，是因为它们涉及的对象真的是以我们描述的方式相互联系着吗？”这样的问题，其答案不能从问题本身上寻找，理由是数学问题具有如下的客观特性：对它的解释取决于我们有多大的控制权。举例来说，前面我提到过，存在性陈述的证明可能远非独一无二的，原因很简单，可能有许多对象都具有所需的特性。但这只是一种简单的数学现象。你可以接受我的分析，相信问题中的对象是一种“真实的存在”；也可以将这一陈述看成其存在如同游戏中图标在纸面上移动，或是把对象看作通俗小说。事实上，数学的某些部分是不可预料的，而另一些则不同；有些的解是唯一的，另一些则有多个解；某些证明是显而易见的，而有些则需投入大量的精力。所有这些都取决于我们如何描述数学问题的产生过程，都完全独立于一个人的哲学立场。

评蒂莫西·高尔斯的“数学是一种发现还是一种发明？”

吉迪恩·罗森

从某种意义上说，蒂莫西·高尔斯与其说是杰出的数学家，不如说是具有20世纪50年代风格的大众语言哲学家。他在本书的首篇文章——“数学是一种发现还是一种发明？”中，非常仔细地道出了数学家（以及Google数据库中不同程度被提及的芸芸众生）在涉及数学各领域研究时是如何实际运用这些词汇的。高尔斯的结论（大致）是，对于数学家在他从事的工作中难有选择余地的情形，我们倾向于用“发现”来看待这一研究过程；而当手头工作有许多方法可供选择，数学家对他所进行的研究有一定的控制力的时候，我们倾向于用“发明”或是“构建”来形容其研究过程。

高尔斯坚持认为，让他感兴趣的这种区别与一个人在数学形而上学的问题上持什么样的观点无关。

> 你可以接受我的分析，相信问题中的对象是一种“真实的存在”；也可以将这一陈述看成其存在如同游戏中图标在纸面上移动，或是把对象看作通俗小说。事实上，数学的某些部分是不可预料的，而另一些则不同；有些解是唯一的，另一些则有多个解；某些证明是显而易见的，而有些则需投入大量的精力。所有这些都取决于我们如何描述数学问题的产生过程，都完全独立于一个人的哲学立场。

我感到这说的是没错，但它也提出了一个高尔斯没有解决的问题。高尔斯描述了数学家们倾向于说所做出的成绩可归于发现或发明的条件，以及某些成就可能被当作发现而另一些被当作会发明的条件。但我们应该如何认真看待这些语言上和心理上的观察呢？正如哲学家们常常指出的那样，营造一个我们乐于说这说那的氛围是一回事儿，而要确定一种能正确地说这说那的条件则是另一回事儿。因此，我们假定数学家们对一些研究成果划归（譬如说）为发明达成一致，但这是否就意味着，甚至暗示着，这项成果事实上就是发明，还是它不过是一种有可能是错的但被太当回事儿的说话方式？

我相信，这个问题在不同的情况下有不同的答案。正如高尔斯指出的，我们不但谈论很多东西的发明/发现：理论、定理、证明和证明技术等，而且也谈论各种各样的数学对象（数、数系）。我们会说康托尔发明了超限数理论，但我们不太可能说康托尔发明了超限数。我们先重点讨论一下发明/构建这类修辞在数学对象上的应用。在本文里，高尔斯讨论了大魔群的情形。大魔群是一种元素众多的有限群，其存在性和唯一性分别于1982年和1990年被证明。语言学证据表明，数学家们更倾向于说大魔群的“构造”而不是大魔群的“发现”，高尔斯解释了这一点。对大魔群的存在性的证明不是唯一的，我们可以有很多例子来确立其存在性定理，即具有相关性质的群是存在的，尽管（正如所发生的那样）这些例子彼此间同构。但是我们有什么理由要认真对待这种构建的意象呢？在我看来，这是这个短语按字面意义运用时所表现出的一种不可通融的性质：如果一样东西是被发明或构建出来的，那么在它被发明之前，它并不存在；如果它没有被发明，它也不会存在。相反，如果一样东西是被发现的，那么它在被发现之前必须存在（或至少独立于发现的具体细节）。但是，我觉得高尔斯会同意我这样说，在大魔群尚不存在的1982年之前，谈论大魔群的性质就显得很奇怪。如果是这么回事，那么格里斯最初问自己的问题，“大魔群是否存在？”答案应该是显而易见

的："还不存在，但也许有一天会存在。"但事实上没有人会这样来讨论数学对象。因此我倾向于认为，即使高尔斯关于我们在什么情形下倾向于采用发明或构建等词汇来指称数学对象的条件是正确的，在这方面完全从字面上理解这种语言仍将是一个错误。

另一方面，当我们谈论数学理论——尤其是像微积分这样的大的理论框架时，情况则完全不同。如果有人问（譬如说）在1650年之前是否存在一种强有力的代数方法用于计算曲线围成的面积或曲线在某个给定点的切线，或者说是否存在一种深刻的、能够证实这些技术并显示它们之间关系的理论，答案可能是："还没有，但说不定哪天会有。"此外，如果我们这么说应该也很自然：如果没有人写下这样的理论，那么微积分就不会存在。因此，这种理论似乎与小说、诗歌和哲学论文一样，属于相同的本体论范畴。这样的事情都是抽象的作品：当某个人第一次给予具体的表现方式后，这些抽象的实体才得以存在。在这种情况下，我看不出有什么理由不认真地将发明这一修辞视为对基础形而上学的一种清楚的、字面意义上的解释。

高尔斯不主张这种解释，但我不知道他是否会同意我的如下假设：除非我们准备说，本项发明在其被发明之前不存在，否则我们就应该将数学中的发明（构建、创造等）视为隐喻。我们可以接受高尔斯对条件的解释，在这些条件下，我们倾向于将比喻看作对那些清楚的、形而上的中性真理的解释。

第2章　探索巴别数学图书馆

马库斯·杜·索托伊

我是数学家而不是哲学家。我的工作是证明新的定理，发现我们计数领域中新的真理，创建新的对称性，以及寻找不同数学领域之间的新的联系。

然而我的工作描述中包含了一大堆足以引出一些重要问题的词汇。这些问题包括：数学是什么，它是如何与我们生活的物质世界和精神世界发生联系的，等等。这些语词有“创建”“发现”“证明”和“真理”等，都是些非常动人的语词。每个数学家都会认识到自己在某些时候会考虑，他们刚刚做出的新的数学突破是一种创造还是一种发现。数学是一种客观活动还是一种主观活动？数学对象存在吗？

在我看来，解决这些问题的唯一方法是对我在做数学研究时我所想的东西进行分析。所以，我从我的工作经历中选取一个片段来分析，这有助于我探索其中的一些问题（关于这类发现的更多细节可参见du Sautoy，2009）。

作为数学家，我最自豪的时刻就是构建了一种新的对称元，其子群结构与椭圆曲线模p解的数目计算相关。寻找椭圆曲线的解是数学领域最棘手的问题之一。所谓椭圆曲线，就是满足如$y^2 = x^3 - x$这样的方程（或更一般地，y的二次方等于x的立方）的曲线。克莱数学研究所提出过一个悬赏100万美元奖金的问题，称为伯奇和斯温纳顿—戴尔猜想（Birch and Swinnerton—Dyer Conjecture），目标就是理解什么情况下这些方程有无穷多个解，其中x和y都是分数。

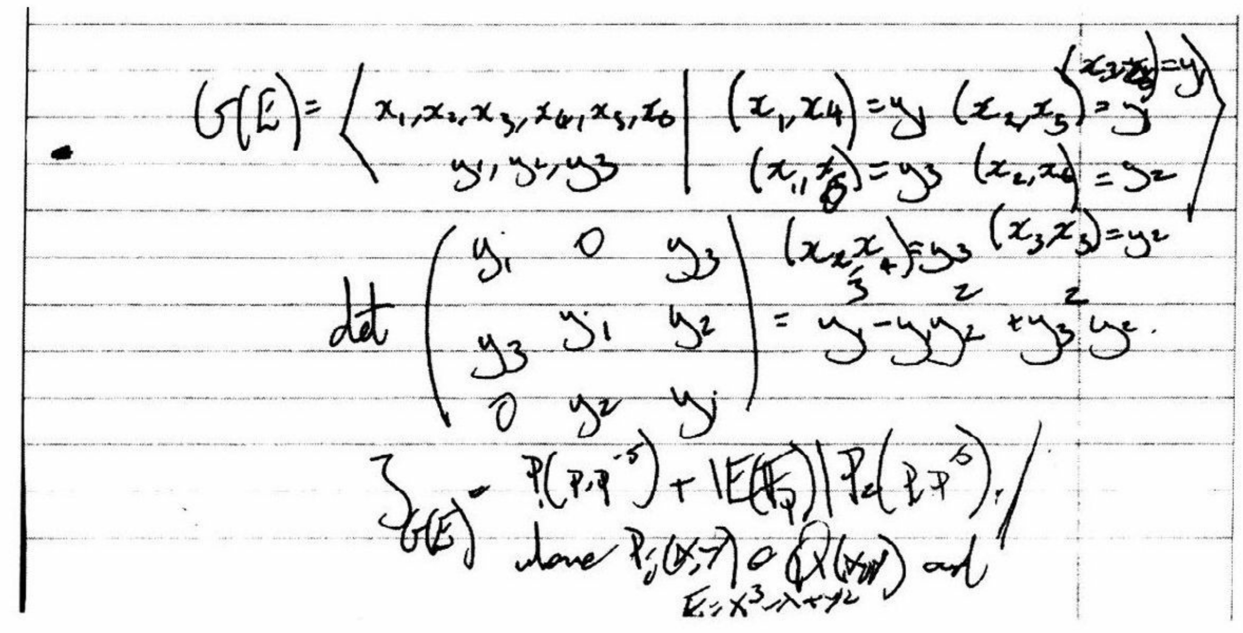


图2.1我的笔记本里记载的一种新的对称元的构建

我构建了这种对称的对象，其结构可以为方程组求解这样的重要问题进行编码，当时波恩的马克斯·普朗克研究所正在进行这方面的工作。我与德国同事证明了的一个数学定理表明，可能存在这种对称元，但在展现这种联系的对称群被构造出来之前，它都可能只是一种错觉。坐在波恩的办公室里的那个晚上称得上是数学家常讲的关键时刻之一，当时我头脑中突然灵感闪现，我急忙在黄色拍纸簿上写下这些新对象彼此间相互作用的对称性结构。这本拍纸簿便是我进行数学沉思的调色板。

感觉很重要。我花了几天时间来真正证明我的想法。一旦细节凸显出来，这个新对象便显示出对称性世界与此前从未显露的算术几何世界之间的联系。

当然，当我说我构造了这个对称的对象时，我不是从物理上构建了它。它是那种只有生活在数学的抽象世界里的心灵方能感知的对象。我既不同于第一个雕刻出有20个三角形面的正二十面体的人，也不同于第一个发现一种新的铺设方式用对称瓷砖铺满格拉纳达的阿罕布拉宫[15]墙面的摩尔人艺术家。我所发现的对象的物理表示只存在于一些高维空间。即便如此，这些表示也仅仅是对基本对称群的表达。正二十面体（参见本书第191页图）和正十二面体的旋转对称性只是这种称为A_5 基

本对称群的两个例子。同样，在阿罕布拉宫发现的这两种设计，虽然物理上非常不同，但在基本对称群下是相同的。

正如数字“3”是由含三个对象（3个苹果或是3只袋鼠）的集合的共性抽象而来一样，对称性632的命名是对阿罕布拉宫的这两面墙所具有的对称性的共性的一种抽象。抽象对称群由每一种对称性的名字来描述。当你一个接一个地研究了对称性后，你就知道怎么去解释这些对称性之间的相互作用关系了。

在波恩的那个晚上，我“构建”的就是这样一种抽象对称的对象，其对称性相互作用的结果是产生出一种与椭圆曲线之间有趣的新联系。它肯定不会在现实世界中存在，然而，如果你在数学世界里花上足够的时间，你就可以得到一种类似于处理正十二面体或在阿罕布拉宫铺墙那样的实体。

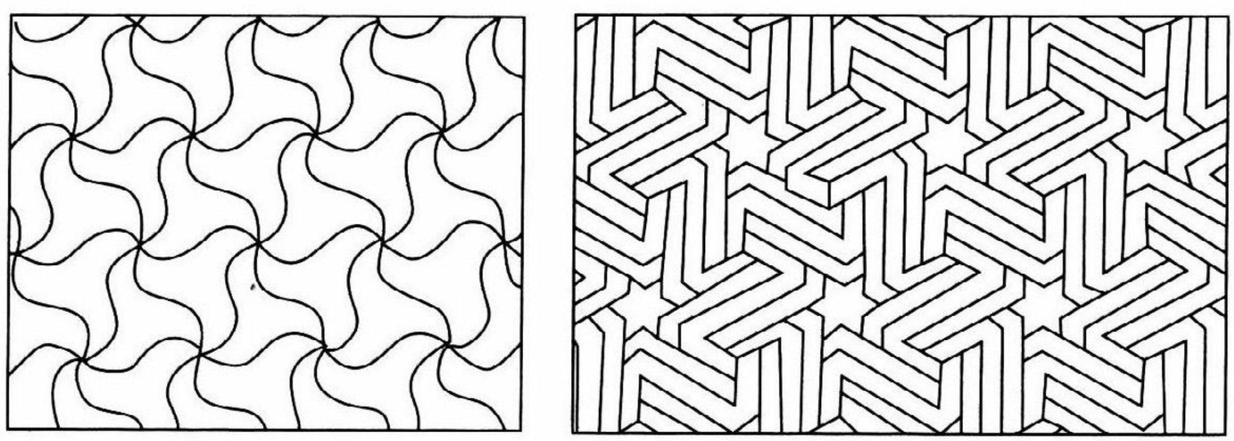


图2.2阿罕布拉宫（Alhambra）两堵墙上铺设的花纹，这种称为632的对称性属于同一个对称群

在描述上述情节时我一直小心地避免使用“创造”这个词。我不得不控制自己别写错字。因为构建这个新的对称群确实像一种创造行为。我体验到一种强烈的责任感——我在黄色拍纸簿上写下的文字正在让某件新东西从无到有，而这件东西在我勾勒出它的轮廓之前并不存在。正是通过我的想象才有了这件东西。它需要借助我这个中介才能来到世上。它不是那种没有我的存在也可以自然演化出来的东西。我给了它生命的动力。

许多数学家都谈论过数学的创造力。正是这种创造力吸引我从事数学而不是其他学科的研究。我觉得其他学科的研究更多的是需要观察。

我在求学的年龄对音乐很感兴趣。我学习小号，喜欢戏剧，喜欢读书。科学不曾真正抓住过我的想象。但在13岁那年，一次数学课后老师把我留了下来：“我想你会明白真正的数学是什么。数学不是我们在课堂上背的乘法口诀表和长长的除法。它要比这些更令人兴奋。我想你一定喜欢看到它更广阔的景象。”他给了我一些书，他认为我会对它们感兴趣，领略到数学世界所展现的各种风采。

其中一本书是G.H.哈代写的《一个数学家的辩白》（1940年初版）[16]。这本书对我有很大的影响。读哈代的书让我明白，数学与创造性的艺术有许多共同之处。它与我喜欢做的事情——语言、音乐、阅读等——似乎是相通的。哈代自己就是这样的一个例子。他是这样来描写数学家的：“数学家和画家或诗人一样，都是模式的创造者。如果说他的模式比其他人的更永久，那是因为他的模式是用概念建构的。”随后他写道：“数学家的模式，像画家的或诗人的模式一样，必须是美的；他的概念，就像颜色或语言，必须以和谐的方式构成。美是这种模式的第一个检验：丑陋的数学在世界上不会持久。”在哈代看来，数学是一种创造性艺术而不是有用的科学。“‘真正的’数学家的‘真正的’数学，譬如像费马、欧拉、高斯、阿贝尔和黎曼的数学，几乎是完全‘无用的’（‘纯’数学的‘应用’真就是这样）。基于数学家工作的‘效用’来评判一个真正的专业数学家的职业生命是不可能的。”

我的对称群性的构建确实不是出于实用的目的。它是体现我的审美意识的一种创造。它是那么令人惊奇，那么出人意料。就像一首乐曲的主题，在证明过程中它突变成一种完全不同的东西。我想，促使我构建这个对称群的一部分动力正是因为它在数学上具有某种实用性。它有可能最终帮助我们更好地理解椭圆曲线。它为我们认识p—群分类的复杂性提供了新的视角。但我仍然认为这种创造性的行为不是那些超出我的控制力范围的外部因素硬强加的结果。

然而……这是不是说数学对象只是待在那里等着别人来发现它呢？是不是我在波恩的那一刻只是一种发现行为呢？也就是说，如果我没发现它，最终也会有别人来得到相同的结果呢？我不过是在数学园地里瞎刨，凑巧发现了这种对称的对象？是不是说它一直在那儿，等着人来揭示？为什么这种发现与科学家首次发现金这种元素或是天文学家首次发现海王星有很大的不同？

关于这些问题，哈代在他关于数学的创造力的一篇演讲中给出了完全不同的思想表达：“我相信，数学实在在我们之外，我们的功用是发现或观察它。我们证明的定理和那些我们夸大其词地当作‘创造’所描述的东西只不过是我们对观察的记录。”这里我妄加归结一下数学家与其成果的关系，我认为所有数学家们都是这么看待他们的工作的：任何创造性的会计运算都不可能使一个素数被整除。正如哈代宣告的：“317是一个素数，不是因为我们这么认为，也不是因为我们的思想经过这样或那样的改造，而是因为它就是这样存在的，是因为数学实在就是这样被构建的。”

发现新的对称群与发现一种新元素或发现一颗新的行星之间也许有一点是不同的：因为黄金和海王星是自然演变的，不需要我们介入。但我还是觉得，如果我没有发现这个对称群，那么一定会有别人将它构建出来。在多大程度上它是我的想象的产物呢？历史上不乏这样的记录：数学对象被不同的数学家以彼此独立的方式同时发现。最有名的当属高斯、鲍耶和罗巴切夫斯基对非欧几何学的发现。虽然他们所使用的符号、说明和解释可能各具特色，但他们发现的对象——一种具有三角形的内角和小于180°性质的几何——是一样的。

相反，人们无法想象三位作曲家会同时创作出《死亡与少女》弦乐四重奏。这个作品是舒伯特天才的杰作，诞生自非欧几何首次面世的同一时期。然而，尽管音乐本身是独一无二的，其他作曲家也永远不可能创作出完全同样的作品，但音乐表现出的这种情绪和变化却完全可以在其他艺术形式上被独立地、同时地表现出来。不同的作曲家经常在相同的时间段里发现新的创作方式、新的曲式结构和新的可能性。舒伯特这曲四重奏标志着音乐创作上浪漫主义时期的开端。但他不是探索狂飙突进[17]时期那种强调变幻不息的键盘曲式概念和强烈对比手法的唯一的一个。我共事过的当代作曲家都谈到过发现某种概念所带来的那种打击，就好像不同的作曲家在创作作品时都发现了同样的新结构、新形式。

也许我可以提出建议来说明我在做数学研究时对创造力的感觉。在波恩的那个晚上，我原本可以在黄色拍纸簿上写下很多不同的对称群。事实上，这样的对称群有无穷多个。我要做的就是写下这些对称群的名称，并定义它们是如何相互作用的……瞧……我已经创建/发现了一个新的群。尽管会出现这些对称群以前是否被构建过的问题，但我更感兴

趣的是为什么那晚我构建了特定的对称群后我会那么兴奋。

我认为将自己比作作曲家和作家是有帮助的。我可以在五线谱上随意写下记谱号，给出不同长度、不同强弱变化的音符，我会谱写一首乐曲。或者，我可以坐在打字机前，打出一连串字母或单词，写出一本书。博尔赫斯的《巴别图书馆》[18] 包含的每一本书都是由25个字符组合写成的，每本书都是410页，每页40行，每行80个字符。当然，这个图书馆里有海量的书籍，准确地说有$25^{1312000}$本。

它们都待在那里等待着某位作者去发现。《远大前程》在查尔斯·狄更斯取下它之前早就已经在那儿了。所谓创造性的行为就是从所有可能的书中抽出这本书来写。我认为，数学研究同样是这种情形，只不过我们经常忽视这一点而已。

我可以无止境地写下新的和原有的定理。我可以建立无限多个新的对称群。我可以通过电脑运用以前陈述的每个语句的逻辑推理规则将这些对称群一个个地复制出来。它们都具有客观的真实性。所有这些在数学上都是真实的陈述。但问题是，就像人们对巴别图书馆里的大多数书籍不感兴趣一样，这些新的定理同样是平庸的，人们不感兴趣。

能产生数学真理的东西还有很多。在数学家看来，艺术就是一种可用于挑选逻辑途径的准则。这里，我认为美感在作出这些选择的过程中扮演着重要角色。我想告诉大家，我发现这种新对称群的原因说起来令人惊讶。它就像小说中的某个时刻，当你认为主角该是这样行事时他却突然变成另一种完全不同的行事风格。

也许数学和其他科学之间的区别就在于我们生活在这样一种自然世界里，它充当着代理角色，挑选出那些具有特定性质的东西，而正是这种特殊性使我们——作为科学家——试图理解为什么它们如此特殊并被挑选出。通常情况下，能给出答案的最终只能是数学。

我很赞同本书其他地方所提出的建议：人们经常是在看到由数学得到的结果比运用它之前更丰富时才体会到它的价值。对称群的定义看起来很简单。人们很难相信会由它导致像大魔群和E_8 这样的奇异对象的发现。

文化和历史背景也会对不同的数学发现的认可并为之激动产生影响。每个21世纪的数学家都会关心黎曼zeta函数是否有一个零点不在临界线上。相比之下，关于是否存在奇完全数的问题虽然也是个令人印象深刻的数学问题，但我不认为21世纪的数学家会关心这个问题。这也就是现在没人真正努力要去证明这一事实的缘故。与此相反，如果这个问题放在古希腊人那里，那将是一个令人激动的发现。对它的证明自然也会产生关于数的令人兴奋的新的数学见解。现在是否真有人关心费马方程$x^n + y^n = z^n$是否有整数解？可以确信，不会有多少定理是建立在“假设费马大定理为真，那么……”基础上的。但为什么数学界一直在追求这个定理的证明呢？是因为它能催化某些惊人设想的发现。

人可能会通过宣称“数学发现的是宇宙的永恒真理”来对数学和艺术创作做出区别。我无法让一个定理为真恰恰是因为我认为它应该是美的。如果黎曼假设被证明是假的，那么我们对素数布局是多么漂亮的信念就将被打破。但对此我们无能为力。黎曼假设可能为真也可能是假的，对此创造性思维无法改变。相反，人们不会谈论《死亡与少女》或《远大前程》的客观真理性。因为从一开始，这些作品就引发了观众的多种反应。歧义是艺术创作的一个重要组成部分。但歧义对于数学家来说则是灾难。数学研究中的创造性行为集中到一点：就是提出黎曼假设是否为真这个问题。关于素数我们可以提很多问题，但为什么这个问题非常重要，同样是因为它提出了素数研究中一个非常特殊的问题。当你第一次了解到素数与黎曼zeta函数的零点之间的联系时，会不禁感到喜出望外。这正是一种不寻常的转换。

数学发现的另一个关键是如何整合发现所跨越的主题。这种整合对于数学价值的判断往往是很重要的。那种从主流数学来看似乎孤立的数学发现，尽管令人惊讶或很美，通常可能不会像与其他主题有联系的数学发现那样受到同样的重视。黎曼假设与数学的其他分支有着如此众多的相互关联，这个事实本身就是数学的价值体现的原因之一。这与互联网相似：一个问题的联系越广泛，它在Google的数学问题排名榜上就越靠前。

也许音乐和文学创作在隔绝状态下反而可以做得更好，虽然经常是人们只在它与此前已有作品的联系中才能真正欣赏这些作品。

人们在证明黎曼假设的研究中提出了一个有趣的问题：证明一个猜

想与构建新的数学对象之间是否有区别。当然，构建用于确立黎曼假设是否正确的证明所涉及的创造性与构建新的非欧几何的创造性是匹配的。但是这个过程确实存在差异。这有点像探险家。黎曼指出在遥远的地方有一座山，而那些试图证明黎曼假设的人则试图在数学园地里找到一些路径以便到达这座山。鲍耶对非欧几何的发现就像一位探索者在海洋中遇到一个以前从未见过的新的岛屿。

“数学对象是否真的存在”这个问题又如何呢？我在内心里当然是一个柏拉图主义者。有一些东西确实独立于我们的存在或我们对其想象而存在那里。素数、单群、椭圆曲线，均是这样的一种存在，而不是哪位数学家制造了它们。但后来也许我正在回到这样一种感觉：我的对称群只是对一直在那里的数学实体的一种表达。我觉得我很赞同克罗内克的说法：“上帝创造了整数，其余都是人的工作。”但这并不是说，黎曼假设的真假也是某个男人或女人提出的。正是要么是真要么是假，才使得素数按黎曼假设所预言的那样分布。但做出攻坚这个数学问题而不是其他假设这一决定的则是我们这些男人或女人。同样，我认为，数学家的作用是论述整数的具体性质，并指明那些真正令其他数学家感兴趣的和令人惊奇的特性。我认为，在挑选大的数学问题时需要用到审美判断，就像音乐家在创作伟大作品时需要审美判断一样。数学的实用性对数学家的研究几乎不起激励作用。数学发现最终被应用到现实世界中往往是几个世纪之后的事情了。相反，数学家被吸引到数学上来，是因为数学充满了优美、典雅和惊喜。在数学24证明里，主题被确立后才有变异、交织和产生令人惊讶的联系的那一刻。在我看来，这些特质既可以创作出令人振奋的音乐，也可以带来令人兴奋的数学论证。

对数学家来说，证明的过程，无论是首次踏出一条前人未知的路径，还是跟随别人的脚步前行，都反映了数学的本质。这种本质不是被证明了的、干干净净的定理的陈述所能反映出来的。例如，费马发现存在这样一个惊人的事实：除以4的余数为1的每一个素数总是可以写成两个平方数之和。譬如41是一个素数，它除以4的余数显然为1。费马大定理保证这个素数可以写成两个平方数的和，对于本例就是25+16或5的平方加4的平方。

在数学家看来，这个定理是令人兴奋的，因为它连接两个不同种类的数：质数和平方数。但数学家的真正快乐在于找到一种证明方法来证明为什么会存在这种联系。当你突然看出素数和平方数之间为什么会有

这种共同性时，你的心头就会为之一震。它们就像是一个共同主题的两种不同的变化。

如同人们开始量化研究是什么构成了优美的音乐（研究者试图绘制出音乐的不同特征）一样，我们有可能借助一些方法来判断为什么我们会给一些证明了的数学定理以大奖，并公之于《数学年鉴》刊物，而对另一些证明则不感兴趣。这跟证明的复杂性有关吗？有时确实有关，虽然简单性往往是数学家的一盏指路明灯。四色问题的证明是复杂的，但不漂亮，因为在你突然明白为什么给地图着色是四色而不是五色就够了的过程中它不能提供那种相当神奇的“啊，哈”一刻。费马大定理的证明相当的复杂（肯定不像费马说的在菜单边页上写不下的那种），使得阅读它的数学家被其中的概念搞得晕头转向，就像一出宏大的瓦格纳歌剧达到大结局的那一幕，但只有这样你才能体味到什么是怀尔斯引领下的不可避免的旅程。第三种测度数学价值的方法是看所证明的结果与其他数学结果之间的统一性，数学的Google评级即是这样的一种方法。

但试图定量刻画什么是好的数学是注定要失败的，就像我们无法用测量来评价为什么莫扎特的音乐是如此神奇一样。哈代在《一个数学家的辩白》里这样写道：“很难定义什么是数学之美，但这种美同其他形式的美一样真实——我们可能不知道一首优美的诗的确切意思是什么，但这并不妨碍我们从阅读中获得美的享受。”

我常常觉得，有关创造/发现的问题与那种关于先天的还是后天培养的争论有共同之处。一个孩子到底多大程度上可归因于基因的遗传呢？在孩子的成长和性格塑造方面环境的影响是否更大呢？数学家发现的定理就是他们的孩子，他们的遗产。一个定理的诞生往往是长期辛勤劳动的结晶。它们的存在是我们继承传统的一种方式。对它们的证明的持久性使我们有机会变得不朽。但这些定理难道就像遗传代码决定了其性状和存在方式一样只是我们工作所用的逻辑框架的结果吗？或者说，我们对所创建的这些定理的培育只是文化——我们身处其中的数学环境——的一种功能？对于那些喜欢非黑即白，要么对要么错，要么证明要么证伪，很少考虑到第三种可能的数学家来说，这不是一个非常令人满意的答复。但也许这就是为什么所有这些哲学沉思进行到最后，数学家往往还是扭头回到自己的数学园地，继续在这片绚丽的风景里跋涉——证明新的定理，构建新的数学结构，陶醉在其不变的确定性中。这就是数学家的工作。

评马库斯·杜·索托伊的“探索巴别数学图书馆”

马克·施泰纳

马库斯·杜·索托伊教授用一种简单但有效的区分调和了数学哲学上的实在论和建构主义的立场。可用数学语言描述的结构是一种独立于我们知识的存在，这是实在论的观点。数学家从这些结构中挑选出那些所谓数学结构。但能够用数学语言描述的未必一定就是数学结构。杜·索托伊教授补充道，美学因素在决定什么是值得探讨的——即所谓数学究竟是什么——方面发挥着主导作用。这也正是我在我的《作为哲学问题的数学适用性》（1998年出版）一书中所采取的立场。我的观点是，从支配数学研究的人类中心论判别标准（一如美学标准）上看，数学是符合人类中心论的立场的。

杜·索托伊以赞许的态度引用的哈代的观点——优美的数学从来不是“有用的”——到底要告诉我们什么呢？我看不出杜·索托伊教授有什么理由接受一种明显错误的观点，可能主要是基于一厢情愿的想法。（哈代不想让数学用于战争。）相反，哈代这么说可以说是大错特错，因为许多科学家相信，数学越优美，它的应用就越广泛。哈代写道：“没人会知道数论或相对论会用于战争目的，似乎不太可能会一做这么多年。”虽然说爱因斯坦发明了原子弹的观点是荒唐可笑的，但另一方面说质量与能量的等价关系决不会用于战争目的同样是荒唐可笑的。至于数论，据我所知，该领域的大部分工作是简单的分类，因为它可以并且已经被用于密码学。如果有人想出了一种很好的大数算法，他很可能会被逮捕。

我向杜·索托伊教授提出这样一个问题：你能解释一下为什么优美的数学往往在应用上都非常有用？

第3章　数学实在

约翰·波金霍尔

数学家从事的是发现，还是他们只是构造一些精巧的智力拼图，为对此有兴趣的大众提供一种消遣和娱乐？数学只是一种让人费解的逻辑上的同义反复，还是说数学能够提供比它所蕴含的平庸判断更有趣、更有意义的东西？

寻求这些问题的答案不只是对数学本身的尊严和重要性进行评估，而是其寻求的结果能够为更广泛、更深入地探讨哲学问题提供一种重要的思想源泉。数学的进步能够为基本的形而上学问题——“究竟什么是实在的意义？”——提供答案。那么它们能扩展到那种在时空舞台上仅通过物质成分间交换能量就能充分描述的领域的前沿吗？在唯物主义者看来，后者才是实在的真实内容，人类谈论的所有其他东西，譬如精神或价值论之类，不过是谈论物质所附带的现象的一种便捷方式。或者反过来说，真正的本体论上的充分性是不是要求比物理主义能够清晰表达得更丰富呢？

数学实体的本性问题为探寻这个普遍性问题提供了一个方便的检验案例。在这个问题上，我们不妨来看看一篇出版的报告。这篇报告是两位著名的法国学者——分子神经生物学家兼坚定的唯物主义者让·皮埃尔·尚热（Jean—Pierre Changeux）与数学家兼坚定的数学实在论信徒阿兰·科纳（Alain Connes）——之间的对话实录（Changeux and Connes，1995）。尚热断言，数学实体“存在于创造它们的数学家的神经元和突触中”（同上，第12页）。而科纳则认为，在数学世界里，存在一种“比我们周围的物质实在更稳定的实在”（同上）。在这场对抗中，两种完全不同的形而上学观点彼此尖锐对立。

形而上学

哲学的一个基本问题是，我们应该如何看待认识论（知识）和本体

论（存在）之间的相互联系。在一个极端，譬如在康德那里，二者被看成是完全分离的。在他看来，我们所能知道的只是现象，即事物的表现；而本体，即那种自在的东西，是我们所无法看见的。而实在论者所在的另一个极端则认为，认识论和本体论是密切相关的，我们所知道的知识应该被视为了解那种情形下事情原委的可靠指南。几乎所有的科学家，自觉或不自觉地，都是实在论者。如果他们不相信科学知识能够告诉我们这个世界实际是什么样子的，我们就很难理解人们在纯科学研究中所付出的艰辛劳动和巨大精力。然而即使是在科学领域，物理学尽管制约着但却并不能完全决定形而上学。在这两者之间没有简单的衍推关系（entailment）。它们的相互关系类似于建筑的基础和最终屹立其上的大厦之间的关系。例如，量子物理学本质上无疑是概率性的，但这种不确定性的出现就一定是因为我们对精细结构的无知造成的吗？或者说，自然界本质上就存在这种内在的不确定性吗？虽然大多数物理学家都追随尼尔斯·玻尔及其后继者采取后一种观点的立场，但戴维·玻姆表明，对量子理论可以有另一种解释，它能给出同样的物理预言，但这种理论却对应于前一种选择——隐性决定论（Bohm and Hiley，1993）。由于两种理论的经验性结果不存在差异，因此我们不可能在严格的物理学基础上对玻尔理论和玻姆理论进行取舍性选择，而必须求助于形而上学，譬如自然性和缺乏手段之类。

在有关数学实在的性质方面同样会出现类似的问题。数学知识极为丰富，令人印象深刻。那么这些知识与实在之间究竟是什么关系呢，如果存在这种实在的话？尚热在为自己的观点辩护时说道，他采用的是“博物学家的立场，不借助任何形而上学的假设”（Changeux and Connes，1995，第213页）。瞧，要想让我们以别人看我们的视角来看待自己有多困难！在实在这个问题上已形成世界观的任何人，无论其所持的观点是狭义的还是宽泛的，都具有明确的形而上学判断，就像他们用散文来表达自己的信念。唯物主义没有例外。秉持还原论立场的物理学家往往有这样一种误解，他们不知出于何种考虑，想当然地认为自己不在必须先验地做出形而上学的假定之列，他们只信奉科学的可靠性。然而事实是，科学以其有限的实用目的性换取了非常巨大的成功，这一点从量子物理学的例子中可见一斑。量子力学带来的各种发现限制了形而上的思考，但要确定其结论应该是什么，仅靠这些发现本身是绝对不够的。把思想还原到神经网络的物理状态来，这不是神经生物学的本身的推论，而是一个建立在该科学领域之上的形而上学的假设。当然，我无意挑战这样一种信念：人类的心灵活动与大脑行为之间存在着联系。

我当然接受人是身心二元的实体的观念，但这种关系的性质却不是一个仅靠神经生物学研究就可以解决的问题，尽管这种研究无疑很重要也很有趣。形而上学的问题要求有形而上的答案，这种答案必须得到形而上的论据支持。

与以前许多世纪的思想观念形成鲜明对比的是，当代社会似乎把唯物主义当作天然默认的立场，几乎不需要任何论据为其辩护。它所呈现的世界图景就像是一幅月球地貌——庞杂、单调重复，各种信息处理系统充斥其间，但就是没有人。那种令人类生命变得宝贵和满意的大部分东西被当作附带出现的现象，那种催生了科学和数学的创造性的想象力没有得到应有的尊重。个人经验——我们在处理实在这类问题时所依赖的最重要的基础，而非被给予的应有的特权——因其重要性受到无端的猜疑而被摒弃了。我们的精神生活——我们所有知识的实际源泉——被看作好像只是物质活动的副产品，通过抽象以一种神奇的方式被直接取代了。

在唯物主义观点看来，人类不过是肉做的电脑。就数学经验而言，这种观点不能令人满意似乎很清楚。数学思维远不是计算的效率所能衡量的，数学见解也不仅仅局限于哥德尔定理所限定的有限公理化系统内。罗杰·彭罗斯特别强调这一点（Penrose，1989）。

数学实在

关于数学实体的理性世界的实在性问题的思考，与人们在反驳唯心主义者对物理世界的实在性的批评所做的类似辩护有一定的可比性。然而，在考虑这些形而上的论证之前，我们首先必须明确一点：什么样的结果可能是所期望的。所得出的结论的特点是要有见地和有说服力，而不是逻辑上的强制性。严格的“证明”语言——所得出的结论只有傻瓜才不会同意的那种证明——在这个语境下是不恰当的，没有人能够强迫一个顽固的怀疑论者放弃他的立场。但这种证明也可能枯燥无味，甚至似是而非，令人难以置信。唯我论者和那些坚持认为世界和我们对它的记忆只在5分钟前有意义的人，其荒谬性在逻辑上都是无懈可击的。摆脱形而上学争论的最好的办法就是声称已经取得了能够取得的最佳解释。

人在面对物理世界时的思考与面对数学世界时的思考这二者之间的第一层可比性与人类感知的一致性有关，或者说，与不同的观察者所报

告的、解释的相互连贯性有关。关于这一点科纳总结道：

> 除了我们的大脑对它的感知之外，什么东西能证明[多强势的一个词！]物质世界的实在性？简言之，我们对（数学实在）的认知与它的恒常性是一致的……因此，与数学实在相关联的只有：用几种不同处理方式给出相同结果的计算，不管这种计算是由一个人来完成，还是由几个人来完成。

（Changeux and Connes，1995，第22页）

可比性的第二个方面是指独立实在具有性质上、层次上的丰富性。科学对物理宇宙的探索是一个不断向更深层次的合理的物质结构和关系进军的过程，这一过程显然是无休止的。随着物质世界逐层展现，它所显示的丰富性越来越强有力地表明，其根源是研究者有限的人类智慧所无法企及的。至于数学，科纳援引哥德尔定理解释道，其含义——算术的丰富性永远不会被包含在有限的公理化系统之内——表明数学实在具有相似的性质特征。他说道，这个定理意味着“包含在关于正整数的所有真命题里的信息量是无限的”。接着他评论道：“试问：难道实在的显著特点不是独立于人类的创造吗？”（同上，第160页）。

可比性的第三个方面与前一个方面有关，指的是探索独立实在时遇到的出其不意性。实在论者关于物理科学解释可以找到强有力支持的一点是，宇宙被揭示出的有关特征总是出乎人们的意料，迫使物理学家不断提出新的概念。如果不是大自然顽固的刺激所施加的无情压力，这些概念可能永远不会被提出。科学努力的结果是发现而不是构建。量子物理学的有悖直觉的许多概念也许是这种现象最明显的例子。如果不是因为光的被观察到的特征这一无情事实的推动，谁会认为波粒二象性这种明显的歧义性会是一种合理的可能呢？在探索数学世界时，我们同样会遇到这种类似的丰富的出乎意料的事情。科纳喜欢列举的一个例子是26个“零星的”有限单群，它们无法被分类纳入像素数阶循环群那样的一般性类别（du Sautoy，2008）。更形象的一个例子是曼德布罗特集的无穷增殖结构，这种结构可能就源于一个看似简单的简明定义。

这种考虑因素有助于解释许多数学家怀有的这样一种信念：他们从事的是对实际存在的实体及其属性的发现，而不仅仅是编个智力游戏那样一种发明，他们不是要沉溺于简单地炫耀他们的技能。杰出的数学分

析大师G.H.哈代在他的《一个数学家的辩白》一书中是这样陈述他的信念的："数学实在外在于我们。我们的作用是发现它或观察它。那些被我们大言不惭地描述为我们的"作品"的定理不过是我们的观察笔记。"（Hardy，1940：1967，第123—124页）当然，我们的物质的大脑作为观测仪器会包含在作出这些观察的记录里，就像它们被包含在我们对周围物理世界的观察里一样。但在这两种情况下，感知手段都不应等同于被感知的实在。尚热企图将数学实体还原成突触存储的事项，这种企图必然会受到类别出错的抵制，这种做法就如同要将一篇文献还原成写就它的油墨和纸一样简单粗暴。

数学研究是一项智力探索，这一观念的合理性因数学思维的直觉性质和在不自觉的创造性活动中所起的作用而得到强化。有些起作用的东西要比能够用计算处理的平庸概念所描述的东西更为深刻。许多青史留名的发现过程表明，强烈的意识参与并不能导致某个深层次问题的解决，而经过一段时间的休整，问题的答案往往会自动浮现在意识里。完成证明的细节基本上只是一种需要延长运用技巧的体力活儿。这方面的一个著名的例子是19世纪的数学家亨利·庞加莱。他曾在将某个问题与福克斯函数理论联系起来的研究上费尽周折，还是没能取得进展。于是庞加莱决定先放下它，出去度假。但就在他启程的那一刻，完整的解决方案自发地跳进他的脑海。他非常肯定自己取得了突破，但他继续去度假，只是在回程时从技术层面上扫清各种遗留的问题。关于存在深奥的数学直觉能力的问题，最令人惊奇的例子也许莫过于哈代的印度同事斯里尼瓦萨·拉马努金（Srinivasa Ramanujan）了。这个自学成才的天才表现出惊人的数学天赋，在数论方面给出了一条由他发现的深刻定理。这一定理的发现不是靠明确合理的证明，而是靠一种直观的、心照不宣的过程悟到的。如果我们将拉马努金的伟大发现看成是具有访问和探索现存的纯理性世界的能力的结果，而不是简单地归结为他的神经组织的某种偶然的精巧结构所致，这肯定更有说服力。

进化

认真看待数学实体的独立实在性的最后一个证据出自这样一个设问：在原始人类的进化过程中，深刻的数学能力是如何产生的？有一点似乎非常清楚：一些非常初级的初等数学的理解力——计数能力、简单的欧氏几何概念和简单的逻辑推理能力——为我们的祖先提供了宝贵的进化优势。但人类是什么时候其能力远远超出了解决日常实用问题的范

围，达到能够提出猜想并最终证明费马大定理，或发现非交换几何的能力的呢？这些能力不仅看上去与传递生存优势没有直接关系，而且似乎也大大超出了这样一种貌似言之成理的可能：大自然使人类从生存的必要性进化出这样一种能力，那是人类的幸运。

进化论解释是否有力关键取决于如何获得合适的环境要素，以及尽可能使好的遗传因子得到遗传。如果原始人类进化所发生的环境就像严格的新达尔文主义的正统假设所假定的那样完全只有物理—生物维度，那么人类的数学能力的出现似乎是一种莫名其妙的多余的东西。然而，我们绝对可以认真地采纳达尔文主义的解释，而无需假设所发生的一切事情都需要得到完全充分的考虑。如果数学实体构成一个独立的实在王国，那么数学原本就一直在“那儿”，甚至在数学家出现之前就是这样。它构成一种智力语境，在这种语境下，该出现的事情最终都会发生。虽然生存压力有利于大脑结构的最初发展，这种脑结构提供了有限的算术和几何思维，但一旦与数学实在有了中等程度的接触，那么进一步的新的发展因素就将开始发挥作用。有助于肉体生存的这种驱动力会通过心理因素的影响（即人们常说的“满意度”）而得到加强（Polkinghorne，2005，第54～55页）。知识带来的喜悦将我们的祖先吸引到探索数学实体的纯理性世界上来，诱使他们不断取得进步，这种进步很快就远远超出了日常实用上的需要。毫无疑问，与此相关的心理感知能力的发展很可能是因为人类大脑具有后成的可塑性，它的大部分复杂结构不是来自基因遗传，而是来自对经验所造成的影响力的响应。相信数学实在的独立性就容易理解为什么我们人类有能力成为数学家，否则人类无偿拥有这样一种能力就会显得莫名其妙。

不可理喻的有效性

如果数学实体是实在的一部分，那么可以预料，它们所在的本体论领域不是一个不与实在的其他领域相联系的孤立领域，而是与其他领域有着微妙的联系。这种情况的一个非常突出的例子是我们对物理理论的理解与数学性质之间所存在的联系。基础物理学在寻求理论发现时，数学是一项非常实用的技术。它使得构建理论的方程具有一种明确无误的数学之美。这种美是一种非常稀缺的审美经验形式，但数学家却可以很容易地辨认出并接受下来。这种美包含了诸如优雅、经济和“深刻”等品质，也就是说，从一个看似简单的起点出发，通过推导，我们可以得到广泛的异常丰富的结果。物理学家寻求优美的方程绝不是纯粹的审美宣

泄，而是出于一种启发式策略的考虑。近代物理学发展的三个世纪的历史一再证明了它的价值。量子理论的奠基人之一保罗·狄拉克终其一生都在追求数学之美，并且非常成功，正是这一点帮助他做出了卓越的发现。他曾这样说过，你的方程是否完美要比它是否与实验一致更重要！当然，狄拉克并不是要说充分的经验事实最终是可有可无的。没有一个科学家会这样认为。如果你求出了新理论的方程的解，却发现这个解与实验结果似乎不一致，这无疑是一个挫折。但它未必就是绝对致命的。毫无疑问，你必须借助于某些近似处理来得到新的解，也许你刚刚得到的解是在一种不恰当的近似下得出的，也许是实验数据有错——我们都知道，这样的事情在物理学里已经不止一次地发生过。因此，希望总是存在的。但如果你的方程很不美观……那只能说没希望了。整个物理学史还没出现过方程丑陋却能畅通无阻的先例。

曾荣获诺贝尔物理学奖的尤金·魏格纳——狄拉克的妹夫——曾将这种数学之美在揭示物理宇宙秘密方面的非凡能力称为“不可理喻的有效性”。这一明显抽象的学科居然能够为我们理解物理世界的结构指明一条道路，它是怎么形成的？为什么数学家在他们的研究中所发现的纯数学的优美形态会如此频繁地出现在我们关于宇宙的结构中呢？当然这里不是从细节上对这个问题展开讨论的地方[我个人认为，它的答案可能要到自然神学里去寻找（Polkinghorne，1998，第1章）]。就我们目前的讨论而言，指出这一事实以及它对于物理学和数学之间深度纠结所蕴含的意义已足矣。很少有人怀疑物理世界的实在性。既然这样，他们就应该准备考虑承认与之纠结的数学世界存在类似的实在性。

数学还与实在的其他方面纠缠在一起。我希望非常认真地考虑人类与审美领域的关系问题。我不认为我们的审美经验只是一种浮在物理学基础之上作为附带现象出现的泡沫，而是我们接近实在的另一种形式。当然，音乐涉及空气的振动，但我们并不能将对它的鉴赏还原到这些振动的傅里叶分析。声波波包对耳鼓膜的冲击为什么会唤起我们的情绪变化，并让我们相信这种情绪体验构成了我们邂逅永恒的美时的有效经验，这仍是个很深奥的秘密。人们经常认为，数学和音乐之间存在亲缘关系，这种关系不仅表现在个体上，譬如经常有某个人在数学和音乐两方面都显示出强烈的爱好和擅长，而且表现在二者的模式具有共性，尤其是在对位音乐的情形下。

另一种审美体验涉及伊斯兰艺术所展示的重要模式。马库斯·杜·索

托伊（du Sautoy，2008，第3章）对格拉纳达的阿罕布拉宫的壁面装饰图案的对称性有过非常精彩的讨论。在19世纪，群论数学家已能证明，能够如此规则地铺满一个平面的不同的基本对称性类型只有17种。而在这座建造于13～14世纪的阿罕布拉宫，所有这17种对称性都用上了。参与设计的伊斯兰艺术家肯定不知道群论，但他们的杰作表明，他们直观地领略到数学实在。

结论

形而上学立场的说服力的评估标准是一件需要认真对待的事情。它足以包含人类在寻求创造智慧的过程中以其敏锐的洞察力所获得的广泛的基本经验。那种简单地通过非法强求一致的战略（在某些人看来不方便考虑便将其切去）所得到的方案肯定不会被采纳。本章的论证是要说明，以应有的严肃性来看待数学的特点和成就的方法不外乎是在承认数学实体的纯理性世界的实在这样一种形而上学的背景下所制定的最佳方法。

评约翰·波金霍尔的“数学实在”

玛丽・伦

为了回应数学是“创造还是发现”这样一个棘手问题，约翰·波金霍尔通过本文论证道，数学家从事的是对数学实体的“理性领域（noetic realm）”性质的研究，因此，数学研究属于发现性质而不是创造性质。波金霍尔将数学的这种研究性质与物理学的研究性质进行了对比。依据这种对比，我们对人类数学活动的解释不必依赖于数学对象的范围，例如，它不依赖于“巧妙的智力拼图”的构建或是对“某个可怕的逻辑上的同义反复进行艰难的解释”。

波金霍尔相当明确地指出：他（抑或其他任何人）对数学基本性质的这种解释不可能用演绎方法来证明。在这里，我们所采用的哲学理论不是经验证据就能够确立的，因此，要想使波金霍尔的理性领域假说胜过物理主义假说，最好采用归纳性质的理由（即那种不建立结论，至多带来可能的理由）。波金霍尔进一步指出，物理主义应当被视为众多形而上学假说中的一种，而不是一种默认的立场，除非我们有确凿的拒绝

它的理由。因此，波金霍尔的策略是从物理主义否定这种理性领域存在的立场出发，反其道而行之来考虑理性领域假说。化各种现象所提供的证据为我所用，并通过这些现象来证明物理主义的错误和数学对象的理性领域存在的合理性。

一个问题只适合用归纳来考虑，而不能用演绎方法来证明，这当然无损于这个问题的意义、价值、重要性或其易处理性质。事实上，当演绎方法在纯数学中蓬勃发展之时，归纳推理则成为实证科学的命脉，几乎每一个重要的理论问题的解决都需要我们超越由经验观察所导出的狭隘的结论才能取得。波金霍尔的讨论采用了实证科学里司空见惯的两种不同类型的归纳推理，以论证他的形而上学结论：一种是类比论证；另一种是最佳解释推理。

类比论证的一般形式为：如果X在a, b，c…方面与Y相同，那么X也（可能）在z上与Y相同。这种形式的推理对于从诸多情形模型中导出真实情形的结论是必不可少的，其中的推论可能相当平凡（城市就像是一张标着标志性建筑、街道和我们能看到相对距离的地图，因此，它也像标有市政厅位置的地图。如果你要去市政厅，你该直行，走到左侧的第三街即是）。更富创新意义的是，牛顿通过行星和球体之间的类比认为，行星围绕太阳的运动应该类似于用弹性弦拴着的小球的圆周运动（弹性弦的“拉”力作用在小球上，使其做圆周运动，太阳作用在行星上的力起着类似的作用）。这种推测性类比需要一些技巧。对于X和Y的任何值，我们能够找到它们的相似之处。关键是要找到足够相关的相似性以保证推出进一步的相似性。

波金霍尔的类比论证（引自阿兰·科纳对让·皮埃尔·尚热激进唯物主义立场的回应），从三个方面探讨了数学探索与物质世界探索的相似之处。他的结论是，数学探索和物质世界的探索一样，存在一个独立的、客观的、真实的客观世界作为它的研究对象。波金霍尔个人觉得，这些方面中的每一个本身就足以引出他的实在论结论，但综合考虑，它们只是加强了整体的类比力度。因此，在对这种类比进行评估时，重要的是要考虑两种探索方式相似的那些方面（即那种不依时间、观察者、丰富性和惊讶程度而转移的“知觉”上的一致性）是否与波金霍尔所断言的（即独立的对象世界的）进一步的相似性存在关联。波金霍尔坚持数学推理的客观性原则，例如，用不同方法彼此独立地推出相同数学结论的不同推理者肯定认为数学家在导出他们希望的结论过程中是不自由的。

但与波金霍尔的进一步断言相关的相似性，即数学探索，会像物理探索那样涉及一个客观、独立的对象领域吗？

在回答这个问题时，我们可以考虑关于数学性质的其他说明是否能够解释波金霍尔提到的现象，甚至优于波金霍尔的理性领域假设。对于数学研究和物理研究之间的相似性，如果确实存在另一种合理且不必借助于理性领域概念的解释，那么由类比得到的原始论据的力量就被削弱了。这时，我们就得考虑波金霍尔论证策略的第二种形式：最佳解释推理。

我们已故的同事彼得·利普顿将这种推理形式描述如下：

> 鉴于我们的数据和我们的背景信仰，我们推断出——如果属实的话——那种提供最佳竞争性解释的东西，这种最佳竞争性解释是我们从那些数据中产生的（只要这种最佳好到足以让我们做出推定）。

利普顿（1991：2004，第56页）

正如利普顿指出的，“最佳”一词在这里需要做些澄清。具体来说，我们可以区分为

> 由证据支持的最佳解释和能够提供最佳理解的解释。总之，最佳解释（可以区分为）最可能的解释和最可爱的解释。

利普顿（1991：2004，第59页）

正如利普顿指出的，倡导最可能解释推理相对来说没有争议，但可悲的是，也相当无用——如果我们有办法知道什么情形是最有可能的，我们一定会推断出相应的结果，而最佳解释推理肯定是将找到哪一种方法更可能作为其目的之一。由于最佳解释推理是一种实用的理论选择规则，因此我们需要对那种能够在提供理解的意义上成为最可爱解释的东西进行说明。这种说明也许可以借助于理论品质（如简单性、非特指性、统一的力量等）来进行。波金霍尔明确指出，无论我们对“最可爱”的程度如何说明，理性领域假设都为他所指出的数学探索和物理探索之间的相似性提供了最可爱的解释。这种解释肯定比尚热的将数学对象还原到现存的“产生它们的数学家的神经元和突触”的解释更可爱——

你还真以为这些神经元和突触会包含数学所探索的那种丰富的、令人惊讶的、普遍可获得的对象？还存在另一些在波金霍尔看来可由理性领域提供最可爱解释的现象？特别是，波金霍尔指出了那些具有突发的和深刻的数学见解的情形（如由庞加莱和拉马努金所陈述的那些情形）。在这些情形里，人类的推理能力已经远远超出了我们可以期待用进化上的优势来解释，或用数学的那种对于实证科学的“不可理喻的有效性”来解释的狭隘的数学范围。就拿最后这一点来说，波金霍尔认为，将数学看作实在的一个维度，使我们能够理解它在发现物理实在方面令人吃惊的有效性，因为人们预期数学领域“与实在的其他维度之间存在微妙的联系”。但是，存在两个对象系统本身并不足以解释为什么一个系统里的事实一定与发现有关另一个系统的事实存在关联：我家的厨房是一个存在，太阳系也是一个存在，但是如果事实证明，我可以从我的厨房的大小可靠地推断出有关太阳系的令人惊叹的事实，我们或许还是会认为这种推理的有效性不可理喻。关于为什么数学领域的存在性应当从其有效性得到合理解释这一点，我们还需要更多的说明。

有很多人是带着怀疑的眼光来看待形而上学的猜想的。他们的理由是，什么样的证据赞成或反对一个特定的形而上学假设常常并不清楚。波金霍尔的这一章的重要优点是，它规定了这次辩论中对手所要求的明晰而准确的术语：为那些不使用理性领域假设的问题中的现象找到一种更好的解释。对于那些（包括我自己）认为可以找到其他解释的人来说，波金霍尔的论文提出了一个严峻的挑战。

答复玛丽·伦

约翰•波金霍尔

我很感谢玛丽·伦，她对我所探求的维护数学实在的观点进行了富于教益的分析。两个对象的存在本身并不意味着相互联系，在这一点上她当然是正确的，但数学与物理之间的关系是一种深刻的，而且显然是内在的联系（不像她的桌子和太阳系之间的联系）。但我坚持认为，这种分析鼓励了认为任何东西都是一个更大的实在的一部分的思想。

第4章　数学、大脑与物理世界

罗杰·彭罗斯

数学有一种独立的实在吗？抑或它只是人类的思想和文化的产物？或许它只是对我们所发现的所谓数学结构那样的东西的一种理想化，而这种数学结构则是对物理世界的结构和动力学行为的一种近似？

在这一章中，我试图阐述数学柏拉图主义在下述两个方面的观点：数学的独立实在的问题和物理行为是否基本取决于这种预先存在的数学的问题。这两个问题构成了本章的基本议题，尽管它们在反对数学柏拉图主义的人那里也许常常被混淆。我将用图4.1来说明我在这些问题上的立场。在图4.1里，我直观地描述了三个“世界”：物理的、精神的和数学的世界，同时也给出了三者之间的神秘联系。[19]

上面提到的数学柏拉图主义的第一个问题就是这里的“奥秘0”——“数学世界”——是否仅仅是我们的精神活动的产物，在此之外不存在任何实在，还是它有它自己的独立存在的实体？如果是后者，原则上我们是否有机会掌握这个世界的全部内容？第二个（独立于前一个问题的）问题——我们用“奥秘1”来描述——涉及数学在物理理论上的作用。数学在我们理解物理世界上毫无疑问是有用的，但这种效用难道仅仅反映了我们在将观测数据整理成某种易于理解的形式方面的熟练程度，揭示了某些方面的物理本质，而所用的数学理论其重要性则在其次，它们仅仅是好用而已？还是像许多理论物理学家所信奉的那样，物理世界的运行对于预先存在的数学秩序真的存在一种深刻而又精确的基本依赖关系？这种数学秩序非常优美和复杂，它们就在那儿等待被发现，而不是我们在探索理解大自然的过程中简单地强加给大自然的。

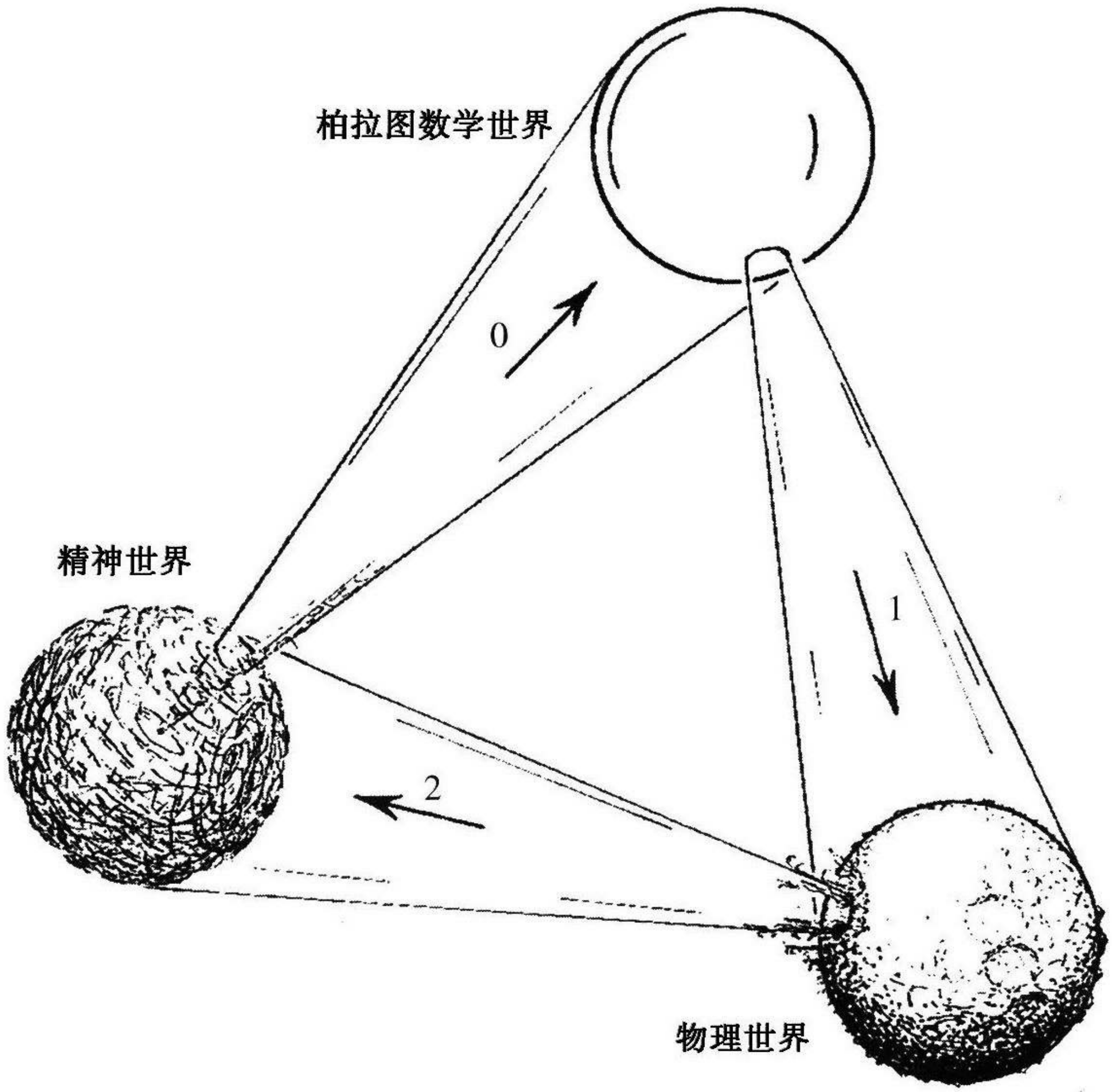


图4.1三个“世界”：物理世界、精神世界和柏拉图数学世界

要使这三个世界构成完整的联系就涉及“奥秘2”，它反映的是物理实在与精神世界尤其是主观意识之间的关系。意识是如何从一个似乎完全由客观的数学运算控制的世界中产生出来的？或者说，在一定程度上看，意识是主要的吗？在任何意义上，它的存在都是我们称之为“宇宙”的这样一种结构存在的必不可少的先决条件吗？意识这种神秘的现象出现在我们的大脑，到底是仅仅反映了我们的脑结构的复杂性，还是它本身就具有某种其他的复杂性质？而且，如果答案是肯定的，那么要理解这种复杂性是不是唯有靠计算？在目前的计算机时代，这似乎已是

一种共识。或者，意识的存在还需要其他的不可能用计算概念来理解的重要的先决条件？如果是后者，那就是说在我们目前用于描述世界的物理学的背后还隐藏着某种东西？或者说我们必须寻找更深层次的（数学？）理论才可能对意识现象进行物理描述？或者说我们可能必须看得更远，对它的理解已经超越了任何一种科学的范畴，采取一种类似宗教在这些问题的基本态度？

在这一章里，鉴于研讨会的主题是数学，因此我主要讨论奥秘0和奥秘1。但是我认为，要对这两个谜团做充分的讨论，就不可能完全抛开奥秘2。我将（借助于哥德尔不完备性）证明，这样一种情形（完全是事实）——我们的大脑能够理解复杂的数学证明，至少在有利的情况下——会使我们得出结论：有意识的大脑的思维不可能完全用计算来诠释，因此，我们的大脑不可能完全是计算物理的产物。不会出现这样一种情形：我们目前理解的物理定律能够包含任何本质上是非计算的东西（这里仅仅具有"随机"性质还不足以被看成是"本质上非计算的"）。由此得出的结论是：有意识的大脑在思维时必然有某些东西是超出我们今天的物理定律所支配的范畴的。我自己的观点是，有意识的大脑思维时的强烈指向性依赖于特定的物理学领域，这个领域可能处于量子/经典的边界，它不在我们今天的物理理论的范畴之内，而所需的物理学革命本身可能还不至于远远超出目前所能理解的范围。

奥秘0的问题确实是一个与上述讨论密切相关的问题。我们将获得数学真理看成是某种"神秘的"事情，其部分原因就在于我们所具有的感知各种特定的数学判断的真理性的能力的性质。正如哥德尔（和图灵）已证明的那样，如果我们将一种具体的、计算上可检验的程序P作为一种有效的数学证明方法加以接受，那么，我们同样必须接受某个命题G（P）的真理性，其中G（P）的真理性不在程序P的范围之内。因此，我们确定数学真理的方法不可能完全还原到我们认为有效的计算过程。虽然不同的逻辑学家对这一结论给出过不同的解释，但在我看来，它的意义很明确：就有意识的理解而言，在纯粹的计算之外必定存在某种必不可少的东西。（有关这一点的进一步讨论，见Penrose，1997。）但在判断数学命题的真理性时有意识的头脑中到底是什么在活动，仍然是个深奥的秘密。

同时，我坚持认为，数学真理具有很强的客观性（事实上，这一点正是哥德尔自己的看法），而不仅是一些在带有任意性的规则的基础上

产生的人类文化“游戏”。不过，我乐意接受这样一种观点：有可能存在“一定程度上的柏拉图主义”，就是说，一些数学家可以将某个命题P的真理性当作一种“客观”事实，而另一些数学家则可能采取P的“真”或“假”是一个见仁见智的问题，主要取决于他所依据的是什么样的“人造的”公理系统的观点。我认为，从哥德尔的建构哲学来看，有一点是明确的：数学的某些领域的确是“客观的”，因此是一种不以人的意志而转移的存在。这个领域可以是所谓“Π_1 句式”的真理，它是这样一类断言：“如此这般的计算过程永远不会终止”（这里的“计算过程”是指“图灵机上的运算”）。Π_1 句式的一个著名的例子是费马大定理。在我看来，Π_1 句式是真还是假完全是客观的，所以Π_1 句式的真值具有一种柏拉图式的实在性，这是毫无疑问的（虽然某些具体的Π_1 句式的实际建立过程有可能含有一定的主观因素）。

另一方面，一些更复杂的断言，如康托尔的（广义）连续统假设，其真理的客观性也许更值得怀疑。所有这种断言其绝对意义上的真假似乎需要一种更强形式的数学柏拉图主义[20]，而不仅仅依赖于某些特定的“人造”的公理系统。数学家是否是一个柏拉图主义者通常指的正是这种意义上的柏拉图主义。我自己的立场并不特别为这个问题所困扰，对于绝大多数与物理相关的论证来说，相对较弱的柏拉图主义似乎完全足够了。

事实上，要接受上述结论，我们需要考虑的是这样一种“柏拉图数学世界”，它仅需大到足以包括对物理定律的描述即可。对于它的“存在”我们有一种额外的、超乎人类文化或想象的情形。就目前已知的情形而言，物理世界的运行在很高的精度上与数学理论符合得相当好。其中特别引人注目的例子是双中子星系统PSR 1913+16（见Hartle，2003），人类观察它已有30年，对它的脉冲信号的时间间隔的观察与爱因斯坦的广义相对论预言之间的一致性可谓惊人（误差小于十万分之一秒）。这表明，自然世界在其最基本层次（这里是指空间和时间的结构）上的运行机制与复杂的数学理论之间有着非凡的一致性。在我看来，认为这一致性仅仅是我们设法让观测事实符合一套我们可以理解的理论框架的结果，这种想法是没有道理的。自然与复杂而优美的数学之间的这种一致性一直就在“那儿”，时间上远远早于人类的出现，或我们所知的宇宙间任何其他有意识的实体的出现。[21]

致谢

本文得到美国国家自然科学基金PHY—0090091的支助，特在此致谢。

评罗杰·彭罗斯的“数学、大脑与物理世界”哥德尔定理与柏拉图主义

迈克尔·德特勒夫森

罗杰·彭罗斯的这一章里包含了许多要求和想法。这些议题已得到广泛讨论。在本篇评述中，我集中评述以下两点，这也是他关于哥德尔定理的意义的观点的核心所在：

Ⅰ.哥德尔不完全性定理表明，“如果我们接受一种特定的、计算上可检验的程序系统P作为数学证明的有效方法，那么我们同样必须接受某个命题G（P）的真理性，其中G（P）的真理性不在P的程序范围之内。”

Ⅱ.断言Ⅰ.“有明确的蕴含关系：就有意识的理解而言，在纯粹的计算之外必定存在某种必不可少的东西。”

我们可以用更清楚、更熟悉的术语来复述彭罗斯的上述两点：

Ⅰ*.对于任何形式系统P，如果我们接受P的所有公理为真，并且其所有推理法则都是有效的，那么我们就必须合理地接受P的哥德尔句式G（P）[及其P上等价的一致性公式Con（P）]为真。

Ⅱ*.由此清楚地表明，存在一组句子A，我们可以合理地认定，它不可能被形式化（即A不是一个可计算可枚举集合）。

我看不出有什么理由可以信心满满地宣称Ⅰ*或Ⅱ*成立。G（P）及其P上等价的一致性公式Con（P）在逻辑上并不由P隐含，这意味着，在逻辑上必然存在逻辑上不由P隐含的句子。一般来说，我们没有理由相信，所有支持P的推理同样支持这些“额外”的蕴含关系。因此，没有理由认为，对P的合理的接受一定包含着对G（P）/Con（P）的合

理的接受。

对于那些认为能够证明P的证据必定也能够证明确信这种一致性合理的人来说，上述推断可能是错的。不过，重要的是要认识到，证明P的一致性与证明G（P）或Con（P）的一致性并不是一回事儿。构成P的一致性的理由未必对G（P）或Con（P）也成立。对于后者，以下补充命题也是必需的：

补充：如果P是一致的，那么G（P）/Con（P）成立。

然而，这样的理由并不一定包括在有关P的一致性的证据中。

当然，不可否认，对这条补充命题也可以进行论证。同样不可否认的是，P可能有着非常令人信服的理由，这些理由已经包括了G（P）/Con（P）成立的令人信服的理由。它的目的只是说明补充命题的非平凡性质，从而对Ⅰ*提出异议，这表明，G（P）/Con（P）的合理接受性往往是以某种预设形式包含在P的合理接受性中的。

对彭罗斯的观点我们还可以从其他方面提出质疑，但限于篇幅这里就不细论述了。

第5章　数学的理解

彼得·利普顿

我虽然不是数学哲学家，但我对科学解释的本质感兴趣。我希望我们可以对数学理解的本质进行讨论。对于这个问题，专题讨论会上的其他与会者都是专家。如果我们能够将物理解释与数学解释进行比较，在我看来是会有启发的。下面的发言希望能起到抛砖引玉的作用。

我自己在数学解释方面的工作部分是基于以下两个简单而一般方面的考虑。首先，我认为在了解所发生的现象和理解为什么会发生这些现象之间存在鸿沟。从理解的角度看，了解通常只是必要条件而非充分条件。因此，正如大多数人都知道的，月亮总是同一侧朝向地球，但很少有人明白为什么会这样。（一旦你明白了要形成这种现象需要月亮环绕地球的公转周期完全等同于其自转周期，这个问题就迎刃而解了。）我们可以认为，有关解释的哲学的大部分工作都是想回答这样一个问题：如何在知其然与知其所以然之间搭建一座桥梁？实际上，这种鸿沟的存在为那些在如何建立联系问题上自以为充分的答案设置了一道有用的约束，因为它表明，任何将理解上必要的单纯知识当作充分条件加以接受的模式都是不充分的。因此，亨普尔的想法——一个好的解释应能够通过提供相信该现象的理由来提供对该现象的理解——可以预料是不充分的，因为在好些情况下，相信的理由往往只需要简单地知道现象确实发生了即可。

促使我对鸿沟问题的可接受的答案设置另一个约束的第二个考虑因素是解释上的“为什么—回溯（why—regress）”特征。我们中很多人在小时候就发觉这样一种提问法常常令父母惊愕：对一个问题我们得到了一个可接受的答案，但我们会穷追不舍地就这个答案本身继续追问为什么。在我看来，这种回溯的寓意不是说解释是不可能的，而是说，在某种意义上，我们可以用我们不理解的东西来进行解释。B可以解释A，从而为在此情形下为什么会是A提供了一种理解，尽管B本身并没有得到解释。因此，我的学生的计算机崩溃了这一事实可以解释她的论文为

什么迟交，至于计算机崩溃的原因没人知道。这充分表明，理解不是要在解释与被解释的现象之间传递某种“超级知识”。理解似乎并不是某种特殊的认知状态，而是提供某种额外的信息，这些信息本身不需要有特殊的认知状态。

这两点考虑——了解与理解之间的差距和解释上的“为什么—回溯”特征——在物理解释的语境下是没有争议的，但在数学上也许就不是这样。因此，我们似乎可以否认知道某些数学陈述为真与知其为什么为真之间有任何区别。在这里，解释性证明与非解释性证明之间好像存在亲缘关系。在数学情形下，我们也不需要我从“为什么—回溯”得出的寓意，即解释不需要特殊的认知状态。因此，人们可以认为，对数学的理解正是源自数学公理这种特定的认知状态，这些公理就是回溯的终止点。数学上的解释之所以不同于物理上的解释，差别是不是就在于了解与理解之间的差距和解释上的回溯性？

物理解释的两种最流行的模式是因果模型和必然模型。根据前者，解释就是提供关于现象的因果链的信息；而根据后者，现象在某种意义上构成必要条件——它必然发生。这两个模型对于数学解释来说似乎明显不恰当。如果我们选择因果模型，那么似乎就不存在数学解释，因为数学事实，不论它是怎么构成的，似乎都谈不上因果关系，因此没有任何证明是自明的。而如果我们选择必然模型，那么又好像每一个证明都是解释性的，这样，了解和理解之间的差异看起来似乎消失了。就算我们知道某个数学真理不是依据证明而是建立在专家证言的基础上，我们仍知道我们被告知的东西是必然的，因为我们事先就知道，任何数学陈述，如果它为真，那么它一定是真的。

幸运的是，我们还有其他各种解释模型可供选择。明显可得的有某种版本的统一模型。按照这种模型，当我们看明白一个问题是如何被嵌入一个统一的模式时，我们便理解了其所以然。这种模糊的想法可以用不同的方式来阐述，但至少它似乎给出了将解释性证明与非解释性证明区别开来的一种办法或标准。然而，统一模型可能忽略了一些理解数学的方法。通过归谬法来证明就是一种不错的检验方法。对于非解释性证明用这种方法是很自然的，但目前尚不清楚，这些证据不能提供对问题的理解（如果证据确实无效）是不是因为它们不具有统一性。这样说可能更自然些：它们的失败正在于它们没能证明“构成”定理为真的东西，而这种“构造”正是物理因果关系唯一的确定形式。因此，对数学的理

解，也许除了统一的方法论之外，我们还应阐明一种非因果决定论的概念。

到目前为止我能想到的这些评论，主要是数学上针对数学现象的解释，但我还想谈谈如何改进我们对物理现象的数学解释。例如，假设父母注意到，当孩子的行为非常糟糕时，惩罚通常会带来行为的改善，而当孩子的表现非常出色并得到奖励时，随后的行为往往会变得不尽如人意。由此他们推断，惩罚要比奖励更有效。这一推断恐怕不是非常在理，因为对于这些行为模式，我们可以通过观察惩罚和奖励是否完全没有效果来检验，用简单地回归到均值的方法就可以对这些模式进行解释。这里我们似乎喜欢用数学统计结果来对物理现象作出解释。

下面是另一个例子。你向空中扔出一把上面写着很多“英语”的小木条。小木条一边下落一边翻转、旋转。现在，在最低一根木条接触到地面之前你用相机抓拍一张所有木条在空中分布的画面。你会发现，接近水平姿态的木条明显要多于接近垂直姿态的木条。这是为什么呢？这是因为木条呈近水平态的方式要比呈近垂直态的方式多得多这一数学事实。（试想一下，一根中心固定的木条，垂直状态只有两种，而水平状态则可以有无限多种，对于接近水平或垂直的状态，这种不对称性同样存在。）

这些物理现象的数学解释提出了一些有趣的问题。它们是真正的非因果性解释吗？这个数学事实真的算是一种解释吗？关于数学现象的数学解释，我们能从这些物理现象的表观的数学解释里得到些什么呢？

此前我写过（直到最近我才想好）理解恰是解释的另一面：“理解”正是对我们从解释中所得到的东西的一种指称。不过现在，我开始认为这是一种过于严格的理解概念，虽然解释是通向理解的一条途径，但它不是唯一的一条。这一思想的较激进的表述形式是：存在很多种解释所不能提供的理解形式。而不那么激进的形式则是，解释所能提供的理解形式通过其他途径也可以获得。虽然我不想将理解的概念扩展到那么远，以至于使它失去了任何有趣的内容，但这两种形式都让我感兴趣。因此，用科学理论进行工作可以提供一种知其然的知识，它无异于一种理解现象的形式，但这种形式有别于解释所提供的理解。至于不那么激进的思想，就像解释的大多数有益于认知的特性那样，都可以用其他方式取得。

暂且以确定论的思想为例。这种思想认为，解释可以通过证明某种现象必然发生这一点来提供理解。有时我们可以证明其必然性，以至于无需解释就可以提供这种理解。譬如，我们通过体会伽利略精彩的思想实验就可以理解为什么重力加速度与质量无关。假设重的物体加速得要比轻的物体快，那么如果用绳子将重物与轻物绑在一起，从两个物体的质量关系上考虑，轻的物体就会使重的物体的加速度变慢，因此两个物体一起下落的加速度应该小于重的物体单独下落时的加速度。但如果将绑在一起的两个物体看成一个物体，其质量显然要比单独一个物体的质量大，因此两个物体一起下落应比单独一个物体下落的加速度大。但是系统的加速度不可能同时既快又慢，所以物体下落的加速度必定与其质量无关。

由于采用的是归谬法，因此这个思想实验本身似乎并不是一个解释，它也没有提供一条通过对思想实验的解释来增进理解的道路。通过对伽利略论证的分析，我理解了为什么加速度必定与质量无关，但如果你要我解释一下为什么加速度与质量无关，我做不到。我能做的就是给你这个思想实验让你自己去分析。对于物理现象，如果我们能够不经解释就获得理解，那么对于数学的理解是不是也可以这样呢？这种推理似乎有一定道理，虽然伽利略的例子在此并不适用，因为尽管在物理的情形下我们能体会到从知其然到知其所以然的过程，但正如我在前面提到的，这个过程对于纯数学的情形并不适用，因为在这里必然性已经是单纯知识的一部分。

虽然我提出的关于数学理解的诸多问题没有时间在此细谈，而且也超出了这次研讨会的议题范围，但我想另外再提两点来结束我的发言。这两点都来自我自己的工作，它们可能会在物理解释与数学解释的比较方面提供某些启发。首先是解释的相关性问题。有一点我们都耳熟能详且很在理，那就是一个好的解释不是由现象单独能够确定的，它还取决于问这个问题的人的兴趣和知识背景。例如，一个好的答案通常需要给出询问者所不知道的某些知识。

在有关物理现象解释的因果模型的语境下，由于因果链的疏密程度的原因，显然需要考虑与兴趣有关的因素。在每一种现象的背后有无数的原因，但不是所有这些原因都是解释性的。因此，虽然电脑死机能够解释我的学生的论文为什么交迟了，但宇宙大爆炸就不是这样，尽管它是每个事件的因果链的一部分。此外，同样的原因对这个人可能是解释

性的，但对另一个人则并非如此。对此我们可以从文献里举个例子。假设麻痹症的唯一原因是未治愈的梅毒，但大多数带有未治愈梅毒的人都没染上麻痹症。一个人可能会发现这个解释是成立的，因为他被告知史密斯患有麻痹症是因为他带有未治愈的梅毒，而另一个人则可能完全拒绝这种解释。

与兴趣有关的某些方面可以通过对“为什么”一类问题给予更多的结构来很自然地分析。对于许多“为什么”类的问题不宜采取简单的“为什么P？”的形式，而是采取“为什么P而不是Q？”这样的有对比的形式。作为陪衬的Q的选择造成差异，这样不同兴趣的人就可以选择不同的Q。因此，如果实际问题是为什么是史密斯而不是琼斯患上麻痹症，这里琼斯并没有染上梅毒，那么援引史密斯的梅毒症就会是解释性的；但如果问题是为什么是史密斯而不是多伊患上麻痹症，这里多伊和史密斯一样也患有梅毒，那么上述答案就不具有解释性了。（因为你的对话者会反驳：“但多伊也患有梅毒。”）在我看来，大多这类对比性问题会产生一种三角关系，它标志着解释性与非解释性原因之间的区别。粗略地说，在这些情形下，解释所需要的是P的使P与Q之间“产生差异”的原因是什么。P的这个原因在Q上看不到。史密斯的梅毒解释了为什么是他而不是琼斯得了麻痹症，但它不能解释为什么是他而不是多伊得了麻痹症。因为琼斯没患梅毒而多伊与史密斯一样患有梅毒。这套句型是否可以移植到数学解释上来呢？我们可以推测说数学上也有各种形式的兴趣相关性。但是否有场合用于对比分析呢？如果是的话，我们又如何选择帮助区分解释性和非解释性的信息的对比对象呢？

最后，我还要提一下我对“最佳解释推理”这个问题的兴趣。有人认为科学家（和普通民众）似乎经常用解释性思考来引导推断。他们之所以推断认为某个假设是正确的，是因为尽管在逻辑上它不是唯一与证据一致的假设，但如果它是对的，那么它便能为此证据提供最佳解释。如果我们想表达这种思想，那么我们需要做的事情之一就是对“最佳解释推理”这句口号里何谓“最佳”做深入研究。例如，我们是否应该将“最佳”解释理解为最可能的解释？或者说应该把它理解为“最可爱”的解释，就是说，如果它正确的话，那么它便提供了最大程度的理解？

最可能的解释似乎是一个显而易见的选择，因为我们都希望我们推断出的结论具有较高的可能性。但我认为，在这种情况下，这是个错误的选择，因为它可能使得解释者的想法几乎是空洞的，我们可以将它还

原成这样一种说法：科学家推断他们所采取的是最可能的假设。最佳解释推理这一想法的最初的吸引力是它会给科学家们的推理实践带来光明，但要说科学家们喜欢采取最可能的假设那是失之偏颇了。如果我们选择“最可爱的”解释，那么我们会得到一种更有趣的推理。我们远不能简单地说科学家都倾向于认为那种“如果它正确，它就能提供最大程度的理解”的解释一定也最有可能是正确的解释，甚至可能是近似正确的。那些发现这种思想脉络很有吸引力的人将面临很艰巨的任务，因为他们现在需要说清楚是什么让一种可能的解释比另一种更可爱。

在数学推理中，最佳解释推理这个概念行得通吗？这似乎不太可能，因为最佳解释推理意味着对非证明性推理给予不完整的解释，而在数学领域，推理至少是演绎的。当你有了一个证明后，谁还需要求助于推理这种软弱无力的概念来得到最佳解释？但是如果我们从发现而非辩护的语境来看问题，事情就不同了。尽管数学研究的启发绝非一种证明性过程，但像最佳解释推理这样的类似概念有可能是适用的。如果我们对数学理解的本质认识得更清楚一点，这个问题我们也许能够回答。

对彼得·利普顿的“数学的理解”的补遗

斯图尔特·夏皮罗

在冈道尔夫堡研讨会上，我们对彼得·利普顿的文章《数学的理解》进行了极为富有成果的讨论。这种讨论是通过提交到研讨会的论文，并结合他和与会者的其他工作来进行的。他敏锐的洞察力和机智，连同他的谦虚和友善的风格，使得讨论变得特别有趣和令人信服。利普顿令人震惊和悲伤的突然谢世，使这一章无法做任何更新或修订。我希望本文为此能够提供一些背景性的工作，以便能够在更广泛的背景下来评价利普顿的贡献。

利普顿是一位科学哲学家，对解释和理解的概念特别有兴趣。他的书《最佳解释推理》（Lipton，1991）现在已快要成为经典——已成为对这一主题感兴趣的任何人的必读书。该书的第二版出版于2004年。这本书——或者说他的大部分工作——的重点，是关于科学的解释和理解。他为研讨会写的这篇文章则是以初步的、有计划的方式，将这些概念和想法扩展到数学上。

正如利普顿在本章中谈到的，知道一个给定的命题为真（或如他所说，知道一个给定的现象发生）与理解为什么该命题为真（或理解为什么会出现这一现象）之间存在清晰、直观的区别。假设在我们问“为什么”之前我们已经知道了“这件事”，那么解释就是“为什么”这类问题的答案。亚里士多德（《物理学》，第3章）写道：“人们不认为他们知道某件事情，除非他们已经掌握了其‘为什么’。”这段论述突出了解释的重要性。亚里士多德指出，我们往往不仅仅满足于知道某些事情是真实的，我们希望对它为什么是真实的有一个解释。

有关科学解释的文献可谓源远流长，至少可以追溯到60年前（例如，见Salmon（1990：2006），这本著作涵盖了前40年的情形。更简洁的处理，见Woodward（2009））。主要富有竞争性的模型有三种，它们分别将解释与科学定律（普适的或统计的）、因果关系和统一性联系在一起。

概言之，就第一种模型而言，所谓一个事件的科学解释是指从科学定律的角度，加上一些初始条件，来对事件的描述进行推理。从这个意义上说，解释一个事件某种意义上就是要证明这个事件的发生是必然的——按照自然法则，它必然发生。而对于第二种模型来说，解释一个事件就是要给出其因果关系的历史链条。利普顿（1991：2004）偏向于这种解释模型，并且从广泛的意义上来理解所构建的因果关系，但他对其他选项持开放态度。第三种模型将科学解释与自然的统一性联系起来。按照这种图像，适当的解释服务于将各种表观上不同的现象统一起来这一目的。解释应能够证明它们是如何从一个共同的来源导出的。

当涉及数学解释时，问题可以分为两类。一类是数学内部的解释。知道某个数学命题为真与理解为什么它为真之间的直观区别似乎在于该命题在数学内部是如何起作用的，这与科学中的情形一样。某些证明在一些人看来是解释性的，但对于另一些人则不然。在这里，正如在自然科学和在日常生活中所表现的一样，亚里士多德的观点是站得住脚的。

另一类问题涉及这样一种情形：数学事实被援引用作对非数学的事件进行解释。利普顿举了两个例子，都是统计上的。正如他在其他地方所强调的，我们总是被逼着要说清楚待解释的情形究竟是什么，或者如他所说的那样，用于对比的类是什么。我们总想解释为什么一个给定的模式往往会出现：譬如为什么扔出去的木条在下落过程中倾向于水平而

不是竖直下落，为什么孩子在受到奖励和处罚后往往会有某种表现，等等。对此情形，他的观点是，模式的解释往往涉及数学定理，如回归到均值（譬如对木条下落的情形我们有测量理论给出的统计事实）；或在其他情况下运用中心极限定理——现代统计理论的主要内容。在每一种情况下，定理均表明，在有关概率的背景假设下，问题中的模式是可能的。

如果待解释的是单个事件，那么情况就要复杂一些。为什么这次大多数木条棒呈水平落向地面？至少在原则上，我们可以用木条下落的初始速度、当时的空气阻力等因素来解释，而不必借助于利普顿引用的几何事实或有关概率的任何考虑。但更务实、更富于启发的做法是注意到我们身边就有现成的例子，根据有关概率和数学的一些基本假设，这种模式就是容易发生。

对于数学事实被引用来说明非数学事件的情形，我们还可以举出其他非统计方面的例子。这里就有一个不费脑筋的例子：假设我们给一个孩子一堆矩形积木块，所有的积木彼此全同，然后要求她用积木搭出一个矩形方格。她尝试了多次之后还是失败了，并开始疑惑为什么她不能完成这个任务。解释是：给她的积木数量是奇数。还有个难一点的例子。我们都知道雨水总是呈雨滴的形式下落，这是为什么呢？这里的解释就不很直接了，涉及表面张力的概念，以及在表面积不变的条件下球形具有最大体积的数学事实。

但将前两种模式运用到数学上似乎并不合适。我们当然可以谈数学“法则”，如交换律和余弦定理，但这么说可能仅仅是一种说话方式。在科学和日常会话中，对于法则（如万有引力定律）和纯属偶然的概括（譬如在本书出版之前所有的美国总统都是男性）通常是有明确区分的。但这种区分对于数学似乎并不有效，因为，正如利普顿指出的，每一条数学真理都是必然的。更重要的是，数学“法则”似乎并不起着区分解释性证明和非解释性证明的作用。一般认为，数学中的每一项陈述都由“法则”推导而来。数学“法则”的概念在引用数学事实解释物理现象时似乎也不发挥太大作用。同样，因果关系的概念在数学也难有用武之地。说一个数学命题是另一个数学命题或非数学命题的“原因”是没有意义的。

因此，在科学解释的标准模型中，只有统一模型似乎适用于数学。

该模型的主要倡导者之一菲利普·基奇（例如，Kitcher，1989）明确注意到这一点。马克·施泰纳（1978，1980）写过大量关于这两类数学解释的文章，并发展了一个用于统一模型模糊边界的解释（但非常不同于基奇的解释）。这在研讨会上已进行了广泛讨论。

如前所述，在他的《最佳解释推理》一书里，利普顿偏好解释的因果关系模型。在研讨会上，他初步建议，因果关系可能只是依赖关系家族中的一员，解释的概念可能与依赖关系有广义的关联。这一建议有可能将数学带入层叠。这个概念是指数学命题是以某种客观的相互依赖关系存在的。如果是这样的话，那么我们可以认为，对一个数学命题的解释是由它所依赖的那些命题的证明构成的。一个解释性证明也许就是揭示证明的前提与结论之间依赖关系的陈述，而一个非解释性的证明则不需要通过对命题的依赖关系就可以显示其结论为真。

我们不妨走得更远点。假设我们对某些物理事件的理解在某种程度上依赖于数学命题，譬如上述例子中就涉及统计学的模式性质（平铺的可能性）和定理（有关素数的定理），雨滴则涉及几何定理。如果这些关系是真正的依赖关系，那么我们就可以将物理现象的数学解释也带进层叠中去。

吉迪恩·罗森在研讨会上的文章里也援引了命题之间的客观依赖性关系，特别是包括不同领域（包括数学）命题之间的相互依赖性关系，虽然罗森的旨趣是在形而上学方面而非解释的理解方面。罗森曾和利普顿以及我们中的其他人就某些哲学问题的各种依赖性关系的性质、其客观性及其相关的其他问题有过广泛的交流。

罗森提醒我们，著名的逻辑学家和语言哲学家格特勒布·弗雷格（Gottlob Frege）也援引过数学上的客观依赖性关系（例如，Frege，1884：1960）。弗雷格经常用认识论方面的术语来指称这种关系，譬如说“辩护（justification）”“证明（proof）”，有时还用命题的“基础（ground）”这样的字眼，但他的修辞似乎不在于目前关注的焦点上。他将依赖性关系看成是客观的：

> ……这里我们关心的不是（数的运算法则）被发现的方式，而是它们的证明所赖以成立的基础，用莱布尼茨的话来说就是，“这里的问题不是要叙述我们的发现的一段历史，历史的东西在不同的

人看来是不同的；而是在谈各种真理之间的联系和自然秩序，这对所有人始终都是相同的。”

（Frege（1884:1960，§17）；

莱布尼茨，《人类理智新论》，Ⅳ，§9）

弗雷格的依赖性关系也是不对称的：不主张命题基础本身；如果一个命题是另一个命题成立的基础，那么第二个命题就不是第一个命题成立的基础。他试图证明算术和分析是逻辑的一部分，理由就是这些学科的基本命题依赖于一般的逻辑规律和定义。

如果我们认为弗雷格一直在寻找的就是算术的和分析的命题的合理解释，那么我们可以把弗雷格要实现的主张放在利普顿和罗森建议的模式中来考察。弗雷格对知道某个命题为真这一点并不满意。怀疑论和虚构主义暂且不论（见玛丽·伦写的那一章），毫无疑问，我们都知道这些命题。弗雷格试图弄明白的是这些命题为什么为真。而这个问题的回答有赖于算术和分析命题所赖以成立的命题。他的建议是，他提供的有关算术和分析的基本原理的推导是解释性的证明。

正像书的标题《最佳解释推理》所指明的，利普顿的思想所特别关注的一个问题是最佳解释推理。这个思想是，科学家之所以往往会推断出一个给定的命题是真的，或可能是真的，只是因为它对于某个现象的解释是最佳的，或像利普顿所说的那样，是“最可爱的”。他用数学并不采用最佳解释推理的建议来结束他的这一章：

> 在数学推理中，最佳解释推理这个概念行得通吗？这似乎不太可能，因为最佳解释推理意味着对非证明性推理给予不完整的解释。而在数学领域，推理至少是演绎的。当你有了一个证明后，谁还需要求助于推理这种软弱无力的概念来得到最佳解释？（第54页）

随后，他简要地建议我们将思考从辩护的语境转移到发现的语境。

然而，就前述所勾勒的广泛的弗雷格图像而言，为了发现最佳解释推理在数学中的作用，我们可能不必离开辩护的语境。利普顿—罗森—弗雷格的意思是，一些数学命题建立在或依赖于其他一些数学命题。正

如弗雷格指出的，回溯不可能永远持续下去。有些命题本身就是数学大厦的基础，不可能再用其他东西来解释，它们是公理。在文章的前半部分，利普顿表明，当涉及专有的公理时，不存在合法的“为什么问题”，或者至少是不存在合法的“为什么问题”的答案：“因此，人们可以认为，对数学的理解正是源自数学公理这种特定的认知状态，这些公理就是回溯的终止点。”

但是，公理是如何被知晓的呢？传统的观点是，公理是“自明”的。对公理的全面、完整的理解立即引起对它的辩护的问题。不过在我看来，当涉及现代数学时，这种传统的基础主义观点是无法真正站得住脚的。作为公理而呈现的一些命题并不显然，夸张点说，要全面和完整地理解它们，就要证明它们（见Shapiro，2009）。或许我们可以说，至少有些公理之所以被选择出来，并不是因为它们具有任何内在的或自明的证据，而是因为它们可以带来好的，或就像利普顿说的那样，可爱的对定理的解释。这是一种整体论的图像。在数学的良好系统化分支里，定理被证明依赖于公理。用亚里士多德的术语来说就是，定理的“为什么”最终取决于公理。当我们转到公理上来，并问它们为什么为真时，或至少是问我们怎么知道就是它们时，答案是，它们提供了对定理的最佳解释。

第6章　数学中的创造和发现

玛丽·伦

数学哲学的一个重要作用就是对数学领域的现象做出解释，就是说，要对研究数学是一种什么样的感觉作出说明。这种现象的一个方面是数学家们常常认为他们是在发现数学实在的性质，而不是创造或发明这种性质。鉴于数学实践的这方面原因，一种自然的假设是：数学家是在研究数学实在，这种实在既独立于人类的创造性的决定，也独立于我们关于实在的信念。这一点与物理学家研究物理实在有异曲同工之妙，物理实在的本性也不依赖于我们。如果我们接受这个假设和基于它的类比，那么问题就来了："这个数学实在的本性是什么？我们是如何可能获得有关它的知识的？"

认真对待与物理科学的类比带来这样一种暗示：数学家研究数学对象，探究的是数、集合等的性质，就如同物理学家研究原子的性质一样。但是，如果存在一种独立存在的、非物质对象的、远离并高于我们自身所居住的物理时空领域的数学领域，那么数学知识如何可能的问题就变得迫切了。我们关于物理领域的知识源于我们作为肉身的生命与物理对象之间的互动，但数学知识的获取能用这种类比来理解吗？数学家G.H.哈代将数学发现描述为对数学实在的观察：

> 我相信，数学实在在我们之外，我们的作用就是去发现它或观察它。我们所证明的，以及我们大言不惭地描述为我们所"创造"的定理，其实只是我们的观察记录。

哈代（Hardy，1940）

但这种将数学发现看成是对数学领域的观察的解释站得住脚吗？

当然，柏拉图认为，数学领域的知识可以被解释为直接"观察"的结果。根据柏拉图的观点，数学对象属于永恒的形式领域，它由不朽的、

无形的灵魂直接感知。这种感知在作为血肉之躯的人“出生”之前就有了。作为灵魂的物质化身的我们所具有的数学知识是我们通过对这种形式的直接经验的回忆而获得的（柏拉图，《美诺篇》，81 d～86 c）。我们通过数学探究所获得的定理使我们能够记得我们先前对该数学领域直接观察所获得的结果。然而，很多人会发现，关于数学实在知识的这种解释很难让人接受，它需要某些关于心身的强有力但含潜在问题的假设作为前提。也许在最佳解释推理的基础上可以对类似柏拉图上述图像的某种观点进行辩护，因为它尽管听起来可能令人难以置信，但最终必定被接受为说明数学发现这一现象的唯一好的办法。然而，在选择这种解决方案之前，有必要对所提出问题中的现象究竟指什么进行研究，以便考虑哪一种解释是可行的。

究竟是什么使数学家感到像个发现者？哈代所说的发现差不多是关于既定数学理论中定理的发现。但数学家当然也会创建/发现新的数学理论，用以证明定理。因此我们必须考虑，在这些发现中，是否有一种可以被称作是对独立于心灵的、存在于非物理时空的数学对象领域的本性的发现。

有人可能会认为，全新理论的发现为这种数学领域的存在提供了最佳证据。毕竟，某种数学理论的公理或基本前提一旦确立下来，我们对数学定理的发现便是一种由这些前提所导致的发现，或者至少是乍一看，这种“如果……会怎样”（进而“如果数学公理为真，那么什么东西必定为真”）的探究事实上并不需要数学公理以某个基本的数学实在为真为前提（后面还将对此详细讨论）。另一方面，在发展新的数学理论时，数学家往往有意识地为这些理论寻找一些基本假设，并认为这些假设真实描述了数学实在的某个重要方面，而不是以只求结果的态度简单地凭空设立理论假设。数学实践的这方面的现象是不是真的为存在一个独立的数学对象领域提供了有力的证据了呢？

事实上，我认为，从反柏拉图主义者的角度来看，理论发展的现象学实际上要比理论内部数学证明的现象学更容易得到解释。如果对数学发现的现象作出解释需要我们预设一种“实在”来作为我们进行数学判断的基础，那么我认为，这个实在不是数学对象领域，而是逻辑结果所在的客观事实领域。至于我们所关心的数学理论化过程中所呈现的发现意义上的理解问题，有待解释的真正难题（事实上，甚至在我们考虑普通的经验推理时，这个难题已经出现）是维特根斯坦所说的“逻辑必然的

硬度”（Wittgenstein，1953：2001，I，§437），而不是数学对象领域的存在性问题。[22]

尽管如此，还是让我们从理论的发展开始讨论。毋庸置疑，一种新数学理论假设的选择通常远不是随意的，事实上，适当的理论假设的发展往往理所当然地被看作一种显著而极不平凡的成就。但这就要求我们将新数学理论的发展看成对独立存在的数学对象领域的描述吗？

我认为，来自数学实践的证据反对这种观点，认为它对我们的理论发展施加了约束。这些证据不要求我们设定一个有待发现的数学对象领域来解释我们的发现。因为在很多时候，数学理论是作为我们给自己设定的问题的解决方案而得到发展的，在这里，问题的约束条件足以缩小可选解的选择范围（甚至常常只有唯一的解）。我们不妨以W.R.哈密顿发现四元数为例。他告诉我们，“（这个想法）在1843年10月16日开始走入我的生活，或像灵光闪现，并充分发展成熟”（引自Tait，1866，第57页）。哈密顿灵感闪现的瞬间（指他发现公式$i^2=j^2=k^2=-1$，并兴奋地将这个公式刻在布鲁厄姆桥的石墩上）难道是他前世灵魂所思考的真理突然再现？

事实上，正如哈密顿自己描述他15年来孜孜不倦地通过类比二维复数来发展三维四元数的加法和乘法法则时所说的那样，他的瞬间灵感看成在他所设定的问题在给定约束下问题解的突然实现更为恰当。哈密顿的目的是要通过类比二元数组x+iy来为形如x+iy+jz这样的三元数组找到运算法则（其中j是有别于i的-1的平方根）。他给自己设定的约束是满足“等模法则”：两个三元数组乘积的模应等于这两个三元数组各自的模的乘积，即，如果（a+ib+jc）·（x+iy+jz）=u+iv+jw，那么等模法则要求（$a^2+b^2+c^2$）·（$x^2+y^2+z^2$）$=u^2+v^2+w^2$。事实上，如果假定乘法法则仍满足通常的交换律和结合律，这个约束是不可能得到满足的。然而，哈密顿发现，如果我们放弃乘法交换律，那么在某些特定条件下等模法则还是有可能成立的，只要乘积i·j=-j·i且二者均不为零。这一点有效地促使哈密顿沿着这样一条思路去寻找出路：乘积i·j的唯一可能的值是第三个虚数k。

哈密顿最初给自己设定的问题无解，而四元数系的出现成为尽可能多地满足哈密顿原始约束的最佳方式。事实上，如果我们将问题预先设定为将乘法扩展到一个基于-1的n次平方根的数系，同时保留结合律和

等模运算法则，那么这个问题有可能只有三个解——即n=0（实数），n=1（复数）和n=3（四元数）——就是一个逻辑结果的问题，而不是一个数学“事实”。

在比哈密顿的这个事例更平淡无奇的数学理论的发展进程中，我们常常发现，公理化受到这样一种约束的制约：需要提出一套满足业已熟悉的数系或经验系统的基本要求的假设系统。例如，欧几里得提出的几何公理就以满足物理空间中有关点和直线的真理为前提，并通过对如下问题的检验而被“发现”：这些问题被假定用来证明关于点、线和几何形状的性质为真的许多其他结果。除了物理解释，有关数学上熟悉的对象的公理的发展也是普遍的，例如，戴德金—皮亚诺公理的发展过程就是如此。这里，公理化受到要求被公理化的结构是一个ω序列（ω—sequence）的约束（详见本书第162～165页）。事实上，数学上熟悉的对象的公理化往往最初是作为定理出现的，例如，有关我自己偏爱的那种“数学实在”——C*代数——的公理最初就是作为盖尔范德—奈马克（Gelfand—Naimark）定理的一部分出现的，它表明，这些公理定义在希尔伯特空间上有界算子代数B（H）的同构子代数上，其结构已经独立于数学的兴趣。

在所有这些情形里，我们受到使系统被构建的那种东西的约束。而这种约束显然会导致这样一种感觉：在构建形式理论时，我们是“做对了”。在这种情况下，公理会让我们感觉到是“真确的”，而不仅仅是方便的或有趣的，而且即使我们不考虑使公理为真的那些对象的存在性，这种真确性的感觉也是可以理解的。在这些情形下，怎样才算做对一件事，这要取决于我们为自己设立的约束（要求我们给出一种满足不同假设的结构）。我们不必（尽管我们可以这么做）将这些约束看成是由那些被理论断定为真的独立对象所施加。相反，作为一系列最初假设或约束的结果，新的公理化能够以在已确立的理论的语境下定理所起作用的方式来产生。在每一种情形下，真正重要的数学发现似乎都是作为这样一种假设的结果的发现。

让我们回到那种在不接受公认假设（如公理）的背景下所发生的数学理论化的情形。在范式情形下，这样的推理是演绎性的，相当于用公理来证明定理，尽管在此情形下也有运用溯因推理的余地：数学家会这么来推理，假使他们的假设成立，那么这样或那样的结果很可能就是真实的。就眼下由公理推得的演绎证明的情形而言，我们能够考虑从下述

意义上可以得出什么样的结论：数学家常常认为他们从事的是发现而不是创造数学结果的活动。那些从事定理证明的数学家真的发现了由他们的假设所导致的已经确定的结果了吗？或者是否可能存在这样的情形，尽管给人很强的发现感，但他们实际从事的是在公理与某种意义上尚不“存在”的定理之间创造某种联系？如果数学家从事的是发现而不是创造，那么这对于我们关于数学的本质的观点会有什么样的影响呢？特别是，这种发现是一种对独立于心灵的数学对象王国的发现吗？另一方面，如果我们选择将从公理出发的演绎证明看作一种创造而不是发现，那么这种观点能够与数学证明的客观性以及数学推理的适用性的观点不相冲突吗？

一种自然的思路是：是的，演绎数学推理是客观的，它引导我们去发现我们的数学假设的逻辑结果。但是，这种客观性与是否存在一个独立的数学对象王国无关，它完全是逻辑客观性的结果。毕竟，在数学假设A_1，…，A_n 基础上进行的得到定理P的推理，我们不是在证明P为真，而是说，如果A_1，…，A_n，那么P。这种条件断言并未断定数学对象存在。其真实性，我们可以假设，完全取决于下述事实，即P是A_1，…，A_n 的逻辑结果。根据这种认识，关于从公理出发的数学证明的被感知的客观性没有什么特别的地方，它不过是一般演绎推理的客观性的一种特殊情形，一种完全取决于“从……得出……”关系的客观性的情形。此外，正如我们所看到的那样，那种导致发展出新的数学理论的推理也可作类似的理解：虽然这种推理不是从公理出发，但它仍然受到建立在预先形成的数学假设的结果和/或必要条件所确立的逻辑约束的支配。

然而，这种看似很满意的立场是有问题的。一旦我们考虑从A_1，…，A_n 得到P的断言在逻辑上到底是什么意思时，这个问题便产生了。我们当然不是要从中得知，采用一套公认的推理法则可以从A_1，…，A_n 推得P。首先，我们知道，这种分析并不能把握我们通常的逻辑结果的概念：理解二阶皮亚诺算术公理。哥德尔第一不完备性定理告诉我们，对任何一套我们能够制定出的（标准的）推理法则，存在一个（二阶）皮亚诺算术语言的句子G，这个句子逻辑上可从这些公理推得，但它不可用我们所选择的法则推得。但即使不考虑哥德尔定理带来的麻烦，我们也应该在对那种关于其逻辑结论的事实依赖于一套选定的推理法则的逻辑结果进行分析时保持警惕。毕竟，使得一套推理法则成为好

的法则的正是（据推测）因为这些法则尊重逻辑结果的事实，而不是相反。

所有这些简单说来就是：有关基于数学推理客观性的逻辑结果的概念都是关于语义结果的概念，而非句法结果的概念。在语义的相关意义上，P是A_1 ，…，A_n 的逻辑结果当且仅当下述这一点在逻辑上是不可能的：当A_1 ，…，A_n 全为真时P为假。但这种分析只是用另一个逻辑概念（逻辑可能性）来替换一个未定义的逻辑概念（逻辑结果）。如果我们的问题是“什么东西基于逻辑结果的客观性”，那么肯定会在逻辑可能性上出现一个类似的问题：对于一个句子的逻辑可能性或不可能性归结为什么的问题，我们能说什么？

这里正是我们关于数学推理客观性的令人满意的观点的困难所在。可以说，对于逻辑可能性，可用的最佳分析是数学分析：句子P在逻辑上是可能的，如果存在一组理论模型，在其中这句话被解释为一个真理。在这种分析中，P是A_1 ，…，A_n 的逻辑结果当且仅当在所有使A_1 ，…，A_n 为真的模型中，P也为真。如果这种分析是正确的，那么与逻辑结果有关的客观事实的存在性便归结为数学对象（集合理论模型）的存在性。因此，数学发现的客观性终归依赖于数学对象的客观性，我们被带回到如何解释我们关于这些事情的知识的困难上来。[23]

对于逻辑可能性是否存在其他可用的分析方法呢？人们在对待逻辑上可能的具体世界时可能会试图避开抽象的数学对象。但是，即使我们能够以尊重我们对逻辑可能性的直觉这样一种方式来搞懂逻辑上可能的世界里的概念，但如果关于这些世界的事实是建立在数学推理客观性的基础上，那么我们在说明如何能够具有有关“由……得到……”关系的知识方面仍会有困难，因为这个世界被假设为在时空上是与我们自己的世界分离的。由于我们似乎知道关于“由……得到……”的某些事实，因此我们应该对那些基于我们永远无法掌握事实而得到的“由……得到……”的关系保持警觉。

这种担忧可能会导致我们放弃所有试图将逻辑可能性还原到更本质东西的努力。事实上，由格奥尔格·克赖泽尔（Kreisel，1967）、哈特利·菲尔德（Field，1984：1989，1991）的讨论可知，我们应当将逻辑可能性看作一种由相关的形式演绎概念和模型理论的一致性导出的独特

概念。我们可以通过推导和模型来理解逻辑可能性：我们知道，如果（在一个公认的推理系统内）我们从句子S推出一个矛盾，那么S就不是逻辑上可能的；而且，如果（在公认的集合论内）我们可以找到一个模型，其中句子S被解释为一个真理，那么S在逻辑上就是可能的。但是（按克赖泽尔/菲尔德的观点），逻辑可能性是一种由相关的演绎概念和模型理论概念导出的独特概念，因此不应被认为可以还原到这二者之一。相反，菲尔德建议，我们应当将“……在逻辑上是可能的”看成是一元逻辑算符，它和一元逻辑算符“……则不是这种情形”一样不需要“还原”。菲尔德认为，这二者都应从其推理作用的规定得到阐释，而不是通过还原变成更原始的东西。但请注意，这个解释承认，不可约模态事实的存在奠定了数学推理的客观性。尽管这个解释避开了承认存在抽象的数学对象，但它仍然要求我们接受一种支撑起数学发现的实在（虽然是模态事实而不是抽象的数学对象）。对此我们需要再次问一声：是什么允许我们人类能够具有关于这些模态事实的知识？

但对于数学发现的现象学或许还有另一种回应：也许我们可以接受这样一种感觉：我们对逻辑结果的判断具有独立于人的决定的客观基础，但同时认为这种客观性仍是一种假象。这是维特根斯坦在以其约定论立场处理数学时所采取的一种做法。按照维特根斯坦的观点，不管表观上如何呈现，“数学家是发明家而不是发现者”（Wittgenstein，1956：1978，I，第167页）。在证明数学定理的过程中，我们不是去发现数学假设的结果，而是决定是否要将这个被证明了的结论作为新的理论结果接受下来。证明不是要对数学概念的内容进行梳理，

> 证明是要改变我们的语言的语法，改变我们的观念。它造成新的联系，并产生描述这些联系的概念。（它不是要确立它们的存在，在它没造成新的联系之前，这些联系本不存在）。
>
> （Wittgenstein，1956：1978，Ⅲ，第31页）

也许，不管表观上如何呈现，数学里就不存在有关逻辑结果的客观事实，有的只是人所决定的、总能以不同方式呈献的结果？

如果各个数学理论彼此完全独立，从而一个理论做出的“决定”不会影响到另一个理论的决定，那么这种观点也许能够站得住脚。事实上，维特根斯坦自己就认为，跨理论联系本身就有一个如何决定的问题。例

如，将自然数嵌入整数等就有一个选择问题（例如，见Waismann，1979，第34～36页）。但在这里，正如弗里德里希·魏斯曼在他放弃自己的维特根斯坦约定论立场时所指出的那样（见他1982年的论文“发现、创造、发明”），数学发现的现象学强烈反对约定论立场。有太多的例子表明，用某个数学分支下已被证明了的定理可以证明（甚至阐释）其他数学分支所得出的结论。魏斯曼给出了一个有关实数结果在复数域得到解释的例子。且看如下泰勒级数的展开：

$$\mathbf{1/1+x^2 = 1-x^2 + x^4 - x^6 + \ldots}$$

这个展开式对$|x| < 1$收敛，对其他所有实数x发散。但如果我们在复数域上来考虑实变量x，即将上式左边看成复函数$1/(1+z^2)$在实轴上的行为，那么上述展开式就可以解释成复函数在$z=\pm i$处有奇点，由复分析的定理可知，任何幂级数展开式只在以原点为圆心、半径为R的圆内收敛，在其他地方发散。由复函数的这些事实可知，实值函数的表现不可能超乎其外。这样的结果似乎独立于我们的选择，更谈不上是人类的约定。正如魏斯曼所说，这让人感觉到好像实函数事先已经知道存在复数。

与这个问题（我们可以称之为数学在数学领域内的适用性）有关的是数学在非数学问题上的适用性现象。有关数学推理被用于经验预测，而且这些预测被证明是正确的这类事实早已引起人们的充分注意。关于数学适用性的一种观点是，数学适用性的根源在于结构上的相似性：一种数学理论（有时）之所以适用于一类非数学现象，是因为这种非数学实在的结构与数学理论所描述的某种结构存在相似性。但如果在数学推理的每一步中，我们都能自由决定所用的数学理论的哪些对象为真，那么这些自由决定是如何导致准确的预测结果无疑就显得过于神秘了。

因此，这两种现象都与激进的反客观论者对数学推理的解释相矛盾。那么，有关数学发现的现象会告诉我们关于数学实在的东西吗？我认为数学定理不会听命于独立的数学对象王国。但如果我们不接受维特根斯坦的极端约定论的观点，那么最起码我们必须承认，我们的数学发现是建立在关于逻辑结果的客观事实的基础上的。如果我们希望站在克赖泽尔的立场上，认为我们所关心的最终问题“不是数学对象的存在性，而是数学陈述的客观性”（Dummett，1978，第xxviii页），那么我们将不得不承认，有关逻辑结果的事实不可能还原到有关数学模型的事

实。如果我们想理解数学发现，我们就必须从这些事实的客观性可能出现的地方来考虑。

评玛丽·伦的“数学中的创造和发现”一文被感知的客观性

迈克尔·德特勒夫森

对于那些从事数学研究的人来说，将数学研究看成一种发现或观察活动而非创新或创造活动是极其平常的事。哈代和哥德尔就是20世纪数学家中持有这种观点的两位著名人物。

伦与哈代、哥德尔一样，认为发现是我们的数学经验的一个重要部分。不过她认为，最有说服力的“被感知的客观性”是有关逻辑结果的事实的客观性。

在玛丽·伦博士看来，事情之所以没有停滞不前是因为存在如下事实：当代模型理论给出的对逻辑结果的最引人注目的处理以及有关模型的判断，使我们深深陷入将有关抽象对象的知识看成是“被感知的客观性”而非逻辑结果的困难境地。因此，她的结论是，数学的“被感知的客观性”不可能毫无问题地归因于逻辑结果的客观性。尽管如此，她并没有看到一种明显更好的选择。

这可能并未充分展现我们在将逻辑结果的“被感知的客观性”作为数学认识论的基准方面的困难。首先，它没有阐明逻辑结果的不同概念的可能性。如果我们将那些有相关数学经验的直觉主义者和其他种类的建构主义者包括进来，那么在逻辑结果的引人注目的实例有哪些的问题上将不会有共识。事实上，对有些人（例如布劳威尔）来说，在对逻辑结果的判断，无论是直觉主义的还是古典的，对于被直接认作数学思维的那种东西的重要性的问题上，就根本不存在共识。

在结果判断上什么东西被认为是客观的这一点也得到了更多的关注。伦认为是语义事实，但也存在其他的可能性。作为一种数学活动，证明是为了在特定受众中取得特定的反应。这表明，结果判断最终应当反映出证明者判断为预期受众的推理标准的那种东西。不论这些标准是以语义术语的方式来规定，还是以可靠的非语义规定法则的行为举止来规定，这都是一个存在深刻分歧的问题。

最后，“被感知的客观性”的复杂性可能被低估了。哥德尔写过，命题“强迫”我们将它们看成是真的。这就是“被感知的客观性”带来的东西？抑或“强迫”只是“被感知的客观性”的一部分？如果是前者，那么被感知的客观性似乎并没有提供多少有关心灵独立的证据。天生的性格，甚至培训所带来的习惯肯定都会引起被“强迫”的感觉。在《算术基本定律》（Frege，1903：1962，Ⅱ，§142）一书里，弗雷格正确地警告道，不要将这种感情冲动当作真理的迹象。正如他所指出的那样：“人们只需要用一个单词或符号往往已足够，并且会产生这样一种印象——这个恰当的名称代表某件事情。随着时间流逝，这种印象会变得如此强烈，以至于到最后几乎没有人对这件事情有任何疑问。”

如果被感知的客观性只是一种强迫，那么它就不是描述真理或客观性的一种强有力的标识。另一方面，如果这种客观性还有更多的要素，那么我们就需要知道其他要素是什么，它们是如何给出比单纯强迫作用更可靠的心灵独立的证据的。

第7章　发现、发明和实在论：哥德尔和其他人关于概念实在性的观点

迈克尔·德特勒夫森

引言

本章研究这样一个问题：我们在获取数学知识方面是否存在某些可用来支持实在论者所持数学观点的特征？更具体地说，这种观点在推理上反映为是否能将下述论断：

Ⅰ.数学家们普遍相信，他们的推理是发现过程的一部分，而不是单纯的发明，

转换为这样的论断：

Ⅱ.数学实体存在于人的心灵可接近的纯粹理性世界。

为了方便起见，我将这个论断当作论证的原始出发点。

对于Ⅱ中“理性（noetic）”一词需要给予简短的评论。传统上它被用来表示某种类型的领悟（apprehension），理解（noēsis [24] ），其特点是具有独特的“知识”性质。通常它与“感觉”相对照，后者是指广泛的感性认知（也称“体验性”认知）或直观。要想很准确地确立某些方法来将非感性实在的知识经验与物质对象的感性经验进行对比，以观察二者之间相似还是相去甚远，这个问题既饶有兴趣又很困难。困难在于“理解”的深度在何种程度上才算满足。因此这样的比较是我们主要关注的问题之一。

经验和非自愿性：背景

在柏拉图看来，理解的对象是形式。他认为，形式不仅表现为经

验，同时也超越经验。另一方面，经验主义者一般则将知识性领悟与对概念（concept）或观念（idea）的领悟联系在一起。这些领悟，如果合乎一定的法则，便构成由对感官体验的抽象而得到的心理表象。

康德非常强调两类表象——直观和概念——之间的区别。他认为，二者之间的重要区别就在于它们在具有判断能力的行为者的控制上处于何种程度。概念被认为是自发的（见康德《纯粹理性批判》，第76页[25]），或能通过判断者自身的知识能动性而变得存在（同上，第439页；Kant，1781：1990，A639/B667）。而直观则不是这样。但最后，康德要求真正的或合法的概念是一致的（同上，第154页；Kant，1781：1990，A150/B189）。因此创建或生成这些概念的自由是一种有约束的自由。

尽管这样，康德看出了在我们对概念的控制与我们对直观的控制之间存在的重大区别（同上，第52页，第107页；Kant，1781：1990，A19/B33，B132）。他将直观及其关系看成是预先给定的，而不处于我们的心灵自发产生的控制之下。另一方面，概念则可以这样来产生，因此不能确保它们一定能由对象来展现。

> ……即便在我们的判断中没有矛盾，它也依然能够像对象并不造成的那样来联结概念，或者并没有一个不论是先天的还是后天的根据被给予我们来赋予这样一个判断以权利；于是，一个判断不论怎样没有任何内在矛盾，也毕竟可能或者是错误的，或者是没有根据的。
>
> 康德《纯粹理性批判》，
>
> 第154页（Kant，1781：1990，A150/B190）[26]

康德对概念和直观之间的这种不对称性的接受提供了一种与当代关于概念性质的观点相对比的非常有趣的观点。具体而言，它似乎与哥德尔在20世纪40年代、50年代和60年代发表的各种奠基性著作中所提出的关于概念性质的观点（见Gödel，1947：1990）大相径庭。下面我们转向探讨哥德尔的观点。

就表象作为其客观性标识的非自愿性这一点而言，哥德尔与康德的意见是一致的。然而，与康德将我们对数学概念的使用看作本质上是创

造性的或自愿的[27]不同，哥德尔认为它明显是非自愿的。因此，与康德相反，哥德尔认为，

> ……尽管它们远离感官经验，但我们对集合论的对象确实有某种类似知觉的经验，这一点可以从那些公理强迫我们接受其为真的事实中看出。我看不出有什么理由我们为什么对这类知觉的信心，即对数学直觉的信心，要比对感性知觉的信心弱。

Gödel（1947：1990，第268页）

事实上，哥德尔更看重我们对概念领悟的这种非自愿性。他相信，这种非自愿性有一种类似知觉但仍属于非感官性质的特性。他还认为，它产生知识的途径是一种既独立于我们心灵的自愿行为又独立于非自愿倾向的重要途径。

> 我的印象是……柏拉图主义者的观点是唯一站得住脚的。对此，我的意思是，数学描述了非感性实在，其存在既独立于人类心灵的行为也独立于心灵的倾向，它只能被人的心灵知觉到，尽管这种知觉可能非常不完整。

Gödel（1951：1995，第322～323页）

然而，尽管抱有这些实在论的信念，哥德尔还是对约定论的某些观点做出了让步。具体而言，他认为约定论者正确地认识到数学是关于概念而非生理或心理属性的学问（见Gödel，1951：1995，第320页）。他说，他们还正确地相信，从一定意义上说，数学真理的真理性在于其词项（terms）的意义——具体点说，就在于由这些词项所表达的概念（同上）。如果说约定论者在哪儿出了问题，他认为，问题一定出在将这些意义看成是由约定确定的这一点上（同上）。在他看来，这里的真理毋宁说是

> 这些概念构成其自身的客观实在，这种客观实在性不是我们能创建或更改的，我们只可以感知和描述它。

Gödel（1951：1995，第320页）

由此可见，按照哥德尔的观点，数学概念是被发现而不是由约定行

为或其他心理行为和倾向创建的。类似的断言对与这些概念相关的真理同样适用（如上）。在他看来，这些断言的主要依据正是这种非自愿性——与数学概念有关的真理“强迫”我们接受其为真。哥德尔认为，这种非自愿性是它们独立于我们的创造力的信号标志，他认为我们的创造力主要由我们的心理倾向和我们行使自愿的心理行为的各种能力构成。

哥德尔的现象学论证

哥德尔原始论证的修正版强调了他将之当作我们的数学经验的广泛的现象学特征的那种属性——即我们的基本数学知识的非自愿性。与原始版本的论证在一开始就阐述的不同，新版本几乎没有对下述事实（如果这是事实的话）——数学家们普遍相信，他们从事的是发现而不是发明——附加任何意义。

他的论证的核心要素是断言：

1.有关数学概念的命题内容“强迫”我们将其作为真理接受下来，其方式类似于命题内容通过感性经验给我们留下印象的方式。

他似乎还进一步认为，

2.这种强加的特点最好解释为将其看成是人类的类似于知觉经验的结果，人类的存在和特性均独立于我们的（个人的和一般的）心理行为（例如，约定或规定的行为）和心理倾向。[28]，[29]

由此，他建议，我们可以正确地推断出

3.数学概念对我们施加一种非感官认知的影响，它们的存在和性质均独立于我们的心理行为和倾向。

因此，对我们的总体数学经验的最合理的考虑意味着

4.数学信念是关于我们发现而非（例如通过规定或约定的行为）发明或创造的客观存在的事物。

总之，这便是哥德尔的论证。这种论证提供了一种比原始论证更加宽阔的表达。我觉得最有趣、最独特的是它诉诸一种假想的、与我们的

数学判断（或至少是部分）有关的“强迫（forcedness）”现象，以及这样一种观念：将这种现象用作数学概念的客观实在性的证据。在下一节里我将更仔细地考虑哥德尔的上述推理，并试图通过某些方法来说明为什么它在论证数学实体的实在性问题上不同于早先的论证。

“强迫”作为实在的一种标识

哥德尔的语言，特别是他关于数学命题“强迫我们视其为真”（Gödel，1947：1990，第268页）的陈述提出了一种广泛的现象学式的推理。更具体地说，这种推理超越了我们对某些命题的明显性（evidentness）的经验。

哥德尔似乎一直特别关注基本的或原始的数学命题——那种其强迫性似乎不能由其明显隐含在（其他）强迫性命题中的那种东西来解释的命题——的明显性。不管怎么说，这似乎就是他在指出不仅集合论的某些命题强迫我们视其为真，而且集合论的公理也是如此时他心里所想的东西。[30]

哥德尔似乎考虑过，我们在数学明显性方面的某些特定经验对揭示下述问题相当灵敏：至少有一些命题是以一种类似于感觉命题通过感官经验“强加于”我们的方式“强加于”我们的。他似乎一直将这一点看成是数学实在的外部标识，这种标识是我们通过广泛的经验可以接近的。

> 存在——除非我错了——一个由全部数学真理构成的完整世界，我们只能通过我们的智慧来接近它，它的存在就像物理实在世界的存在，二者均独立于我们，二者均由上帝创建。
>
> Gödel（1951：1995，第323页）

当然，这种推理也带来了许多问题。其中一些是：

Q1：外部实在的标识对于我们对感觉命题的强迫性的经验有多可靠，有多少启迪作用？

Q2：基本数学命题的强迫性与作为外部实在标识的基本感觉命题的强迫性之间的类比有多可靠，多广泛？[31]

Q1探究的是称为强迫性的感觉判断性质与引起这个性质的外部源的存在性之间的有证据的联系。对这种联系的肯定似乎需要这样一种信念：感觉判断的强迫性最好被解释成能量的某种转移——具体说来就是能量从感觉主体之外的感觉刺激源转移到感觉主体上。

按照这样一种理解，是不是感觉判断的强迫性现象就成为一种可靠的外部实在的标识了呢？我认为，这里我们需要区分两种类型的这种可靠性。一种是我称之为存在可靠性（*existential reliability* ），或曰与强迫经验的外部信号源有关的可靠性；另一种我称之为属性可靠性（*attributive reliability* ），或曰与带给我们感觉判断的外部源的特征有关的可靠性。

当然，有关存在可靠性和属性可靠性的例外情形，我们可以在有关感官知觉的文献中找到，这里我就不一一细述了。但我可以简单地指出一点，像某些幻觉那样的准感觉经验提供了不少案例，其中强迫性经验不是外部实在的现存的可靠标识。同样，光学错觉的著名案例则提供了属性可靠性的类似关切。

除了这些方面，我还应提到三种其他情形。第一种情形与我们对“强迫”的理解有关。当哥德尔说我们的数学经验包括命题迫使我们接受其为真的经验时，他要表达的是什么意思？一种自然的解释兴许包括以下含义：P迫使我们接受其为真意味着我们形成了一种可陈述为P的信念。

这就提出了一个重要问题——哥德尔关于所谓数学“知觉”的认识。它之所以成为问题是因为感性知觉似乎不支持上述含义。也就是说，我们有一种可陈述为P的感官经验（例如，一条线段比另一条长的经验）似乎并不意味着我们形成了一种可陈述为P的信念。有不止一种著名的幻觉（例如，蓬佐幻觉和米—勒莱尔幻觉）让我们感觉到一条线段比另一条线段长而我们并不相信它真的如此。

就是说，在感性知觉上，我们能够有一种可陈述为P的经验，但我们无法判断或相信P是否为真。感性知觉似乎是这样一种领悟形式：不是它所呈现的所有内容都像其呈现的那样真实。哥德尔的数学“知觉”是不是也同样如此呢？也就是说，数学知觉能够在某种意义上“给出”一种不管怎样都不会强迫我们信其为真的情形吗（哪怕这种方式是人为给定

的）？

我不知道这个问题的答案，但这里我主要关注的不是问题的答案，而是问题本身。它展现了——如果我没有弄错的话——由Q2带来的一种特定的困难，一种我们做任何尝试——例如哥德尔尝试在感性知觉与数学领悟的类知觉形式之间进行类比——都必须面对的困难。

这个困难既不是唯一的也算不上是最严重的困难。我认为，更麻烦的是由内容虚假却强迫我们信其为真的那些感官错觉所带来的问题。这方面的一个著名例证是阿德尔贝特·艾姆斯的“扭曲的房间”。如果我们以非立体的方式来看（例如通过一个窥视孔来看）这个房间，那么它看起来就像一个由矩形窗、一方平整的水平地面以及高度和深度均相等且规整的矩形墙面构成的“立体”房间。

但当我们将几个普通对象（例如一个正常高矮的成年人和一个正常高矮的小孩）安置在房间的两侧并分别对其进行拍照，然后将照片合并成一张图像时，怪事便出现了：（例如）孩子会显得比成人还要高。

当然，事实真相是艾姆斯的房间不是普通的立体房间，尽管其外观看上去像。它的看似矩形的窗户、墙面和地板其实都是梯形的，它的墙面的高度和深度不是各处一致的，地面也不平整。这样一种看起来正常，在熟悉的对象被放置其中仍显正常的幻觉暗示我们，心理倾向能够对知觉产生影响[32]。正如R.L.格里高利所描述的那样：

> 显然，我们对矩形房间是如此习惯，以至于我们在观察对象时想当然地认为显得怪异的是物体（人体）的大小而不是房间的形状。但这种情形实质上是一场赌博——实际情形既可能是二者中有一个怪异，也可能二者都怪异。在这个例子里，大脑押错了赌注，因为实验者事先做了局。事实上……艾姆斯扭曲的房间的最有趣的地方是它蕴含的意义：知觉总是将宝押在可获得的证据上……妻子就不会因这种房间而看她们的丈夫显得怪异——她们看到自己的丈夫是正常人，真正奇怪的是房间的形状……（而）熟悉这种房间的人，特别是触摸过其墙面的人……会逐渐减低对其他对象的扭曲效应，并最终变得看出是什么造成了这种扭曲。

格雷戈里（Gregory，1969，第180～181页，

其中括号为本文作者所加）

因此，从某种意义上说，在某些情况下，感官知觉会将虚假内容“强加于”我们，让我们信其为真。而且，它这样做是通过深层次心理倾向的操作而非仅仅通过所涉物质对象的客观属性来完成的。

任何企图在感官知觉与数学领悟这样一种类知觉的形式之间建立类比的尝试都必须解决好这样一个问题：在数学知觉上是否存在像艾姆斯“扭曲的房间”这样的现象？具体来说，我们必须考虑：（1）虚假内容是否像真实内容一样迫使我们信其为真；（2）强迫现象作为外部对象的存在和特征的标识是否比深层次心理倾向的存在和工作机制更可靠；（3）是否存在一种类似于“扭曲的房间”情形中影响感觉判断的影响数学知觉的“熟悉”或“养成（training）”现象；[33]（4）在可靠性与偏好这种熟悉/养成的信念选择之间是否存在一种系统的关联；（5）是否存在一种可靠的唯象方法用以对由外部对象的影响产生的强迫与源自心理倾向活动而产生的强迫进行区分。

此外，还有其他问题。一个问题是一种态度何时可以被鉴别为确实是接受一项命题内容为真的态度。接受可分为各种不同的类型、程度和方式，接受的对象也千差万别。因此，单纯将我们认定事物的意识当作接受事物的态度似乎并不能给确定强迫行为的确切性质和最终来源提供指导。我们的大部分被迫接受的经验无疑是由于长期复杂的而且主要是各类无形的熏陶所现成的，我们通常根本就不知道其性质和来源。弗雷格对我们的许多判断背后所基于的习惯影响持有一种特别不连贯的观点：[34]

> 人们只需要用一个单词或符号往往已足够，并且会产生这样一种印象——这个恰当的名称代表某件事情。随着时间流逝，这种印象会变得如此强烈，以至于到最后几乎没有人对这件事情有任何疑问。

弗雷格（Frege，1903：1962，第2卷，§142）

自从弗雷格之后，不论是哥德尔还是其他人，都没能充分解决这些难题。虽然这不应该被看成表明它们是无法解决的，但它们激励人们对哥德尔的数学发明/发现问题的似乎真实的理解采取另一种考虑。现在

我就来谈谈这类别样的考虑。

传统上的考虑

哥德尔的观点与关于发现与发明问题的更重要的传统观点之间有共同之处。广义上讲，这个共同的关注点就是数学的客观性。更准确点说，二者共同的关注点表现在对数学上某类主观性的限定上，特别是对那些与可能合法使用“创造”或“发明”概念有关的主观性方面。现在我来简要概述一下历史文献中有关这个问题的较有趣的内容。

古代观点

在古代，数学在何种程度上谈得上是创造或发明是个共同关注的问题。但是，翻阅古代文献，我们很快就发现这里存在着用语上的差异。具体来说就是，我们当今所用的“发现”一词的意思与古人在使用该词的所指是相当不同的。

这一点从西塞罗（公元前106～公元前43）对这个语词的使用便可以看出。他区分了两类用于（包括数学在内）任何思想领域内进行细致、系统研究的方法。一类方法他称为“发现”的方法（Cicero，1894～1903，第459页），另一类方法称为“决定”的方法（同上）。随着时间的推移，这种区分成长为两类艺术之间的传统门类：发明的艺术（发明的艺术、发现的艺术或研究的艺术）和裁决的艺术（裁决的艺术，有时也被称为证明的艺术）。

这种区别在科学上和法学上是很重要的。在数学上，人们通常对研究的发现方法与证明方法采取不同的形式加以区分。前者的结果是新知识的有效发展，尽管也许结果的质量不是最佳；后者的结果是知识的完善——特别是对由证明上并非最佳的发现方法所产生的不完善的知识的完善。[35]

由此可见，传统上“发现”和“发明”是同义词。它们标志着经典两阶段证明过程的第一阶段[经典证明的过程可分为认证阶段（*certifcative stage* ）和论证阶段（*argumentative stage* ）]。这种经典证明模式背后的总体思路可以概括如下：

> 一个命题的真正的科学证明 [36] 同时需要认证和论证。仅有论证本身是不够的，因为它太容易被虚构。论证的去想象化恰恰是认证要完成的工作。因此，真正的证明是经过认证了的论证，真正的科学知识需要真正的证明。

这个模式突出展现了古代几何学家的思想。普罗克洛斯（Proclus，411～485）在其《欧几里得〈几何原本〉卷Ⅰ注释》里给出了一个用上述经典方式来解释命题Ⅰ～Ⅳ的“顺序”的有趣的例子。[37]

> ……我们的几何学家（欧几里得，卷Ⅰ）在给出这些问题（命题Ⅰ～Ⅲ）之后提出了他的第一定理（命题Ⅳ）……因为除非他先前证明了三角形及其构造模式的存在，否则他怎么可能言谈关于它们的基本属性？假如有人……说：“如果两个三角形有这种属性，那么它们一定也有另一些属性。”这岂不是很容易……满足这样一种断言：“是否我们知道一种三角形就可以构造出其他所有三角形呢？”……正是为了防止这些反对意见，《几何原本》的作者给我们提供了三角形的构造……这些命题对于这条定理完全是预备性的……

普罗克洛斯（Proclus，1970，第182～183页）

像古代的其他几何学家一样，普罗克洛斯对认证方法与论证方法的概念的区分与古代对问题研究与定理研究之间的区分是一脉相承的。前者由“解决问题的工作”构成（Proclus，1970，第157页），后者主要是“定理的发现”（出处同上）。

前者的目的是“产生、创制或构建某种意义上先前不存在的东西”（出处同上）。这个目标的完成一般是由“构建”一个图形，或将它置于某个地方，或运用于另一个情形下等措施来确保（出处同上）。总之，这个目标通常是通过构建来完成的。

另一方面，定理研究的目标是“看出、确认并展示一种属性的存在性或不存在性”（出处同上）。换句话说，它要努力做到的是“通过对几何研究对象的属性及其内在性质的展示牢牢把握这些对象”（同上，第157～158页）。

因此，问题研究旨在确立某种“新”东西的存在性（即使以前不存在

的某个对象的存在性得到确立），而定理研究的目的是要确立已经存在的东西的属性。因此，定理研究对象的存在性需要预先得到确认。

普罗克洛斯对命题Ⅰ～Ⅳ的顺序的评注反映出他接受了这个模式。命题Ⅳ是一个定理，其对象有三角形、线段、等长线段等，另一方面，命题Ⅰ～Ⅲ是问题，它们是确立命题Ⅳ内容的存在性所必需的。因此它们对于命题Ⅳ是“完全预备性的”（同上，第183页），即使它们在命题Ⅳ的证明中并没有被采用。[38]

在说命题Ⅰ～Ⅲ是命题Ⅳ的“完全预备性”这句话时，普罗克洛斯似乎已经意味着它们解决了命题Ⅳ及/或其证明的一些广泛持怀疑态度的挑战。例如，命题Ⅰ是对问题“你知道完全可以构建出来吗？”引起的挑战的响应。另一方面，命题Ⅱ和Ⅲ则是对“也许没有任何直线段等长于另一条直线段（或这个问题的其他几何对象具有相同属性）”这一挑战的响应。[39]

总之，普罗克洛斯似乎已经接受了这样的观点：组合式的定理研究是一种不可认证的研究，或是一种缺乏根据的研究。为了避免这一点，构成相关定理得以成立的对象应该先被证明是存在的。普罗克洛斯认为，这一点可以通过给这种根据一个一般性的或构造性的定义来做到（这是最佳的还是唯一的？），它是这样一种定义：它“解释一件事情的起源，就是说，这个事情是如何生成的——这就像一种循环定义，即，它就像一条直线绕一定点运动画成的图”（Hutton，1795～1796：2000，第1卷，第362页）。[40]

近代观点

近代数学家和哲学家持有一种类似的观点，这种观点可用莱布尼茨的如下说明来陈述：

> ……欧几里得提出的圆的概念——即由平面上一条直线绕一定点运动画成的图的概念——提供了一种真实的定义，因为很明显，这样的图是可能的。事先……有（这么个）定义……是有用的……（因为）我们不能放心地给出有关任何概念的证明，除非我们知道它是可能的；对于那些不可能的或涉及矛盾的事情，这些矛盾也是可以被证明的。这就是为什么对任何可能性我们需要真实定义的重

要原因。

莱布尼茨（Leibniz，1683：1973，第12～13页（294）[41]，括号为本文作者所加）

莱布尼茨因此认为，根源性或构造性定义是有价值的，因为它们使其所定义的概念的先验的可能性变得已知（见Leibniz，1764：1916，Bk.Ⅱ，ch.Ⅱ，§18）。他认为，这反过来又以某种方式使证明变得“安全”。

但是，他对那些不可能的概念及其存在性做了各种令人困惑的断言。有时他认为存在不可能的概念（即隐含矛盾的概念）。

一个概念要么是适合的，要么是不适合的。合适的概念是那种一经建立就成为可能的概念，或者是那种不隐含矛盾的概念。

莱布尼茨（1903，第513页）

另一些时候他认为不可能的概念是不可能存在的。因此（例如）他声称，“我们不能够对不可能的东西有任何概念”（Leibniz，1978，第4卷，第424页）。他还坚持认为，“那些实际存在的东西不可能丧失其可能性”（Leibniz，1978，第7卷，第214页）。[42]

尽管如此，我这里主要关注的是：当有关A的证明缺少有关A的可能性的知识时，这种证明便存在不安全因素。具体来说，有几方面理由使我关心莱布尼茨可能曾相信存在这种情形。

他提到的危险是存在矛盾的可能性——“不可能的东西……矛盾可以被证明”（同上）。虽然我们今天来看，这似乎不正确。在莱布尼茨时代，定理的主要类型是“所有的A是B”这样的形式。这种形式下的矛盾可能是“有些A不是B”之类形式的句子（或其等价的句子），这需要知道各种A的存在性。

不过，当A是不可能的事物时，各种A便都不存在。因此，那些像莱布尼茨一样认为“实际存在的必然都是可能的”的人可能不会相信，对于不可能的A，我们可以证明“有些A不是B”。因此，他应该没有理由

将“有些A不是B”的证明当作对在A是不可能的情形下证明“所有的A是B”的一种威胁。

更可能的情况是，莱布尼茨在设想“所有的A是B”这一证明的反证时，想的是要证明“没有A是B”而不是“有些A不是B”。如果确实如此，那么我们也还有这样的问题——即使在这种情况下，造成“不安全”的因素是什么。但不论莱布尼茨所设想的危险是什么，它都不会是一种对虚假陈述的直接证明。这个结论可以从基本的（经典的）逻辑事实得到。当A是不可能的事物时，则必然不存在各种A，此时说“所有的A是B”和“没有A是B”都是对的。[43]

但如果不是虚假性反对我们从A的真实定义那里得到保护，那么我们从什么地方得到保护呢？在这一点上莱布尼茨不是很明确。[44] 他可能已经将这种威胁看成是虚构化的威胁——即进行没有根据的研究，从而在最后的分析中一无所获的威胁。这似乎一直是经典的观点，也曾是普罗克洛斯一直追求的观点。可能我要论证的也正是这一点。莱布尼茨时代（也包括其较早和较晚时代）的其他人在安全性问题上要更加明确些，他们一般都给出下述两种观点中的一种：一种观点大致较实用，另一种更理论化一些。我将描述并讨论这些观点。就目前而言，要牢记的要点是：（1）在数学史上，过去曾有过，也许目前仍存在，一种将可靠性与广泛的体验联系起来的传统，至少在某种程度上建构是被当作经验的媒介的，但是（2）这种经验的内容不一定就是最终被证明的或是合理的命题。

作为实际考虑的真实定义

对真实定义（real definition）的实际辩护并不对通常的观念——一个概念的实在性或可能性是由其一致性构成的——构成挑战。它的要求毋宁说是，确立一个概念的一致性的唯一可行的方法是用一个实例来表现，就是说，给出它的一个真实定义。[45] 真实定义作为一种确保一致性的手段，其有效性通常被视为这就是为什么经典几何（其中真实定义占主导）中的问题要比代数中的少很多的原因。普莱费尔（J.Playfair）的下述论述可谓18世纪对这一思想的一种典型的表达：

> 几何命题从没有引起过争议，也不需要形而上的讨论的支持。另一方面，在代数中，关于负数及其结果的学说则经常困扰着分析

者，将他卷入最复杂的争论中。这种多样性的原因毫无疑问必须到被我们用来表达概念的不同模式中去寻找。在几何里，表示每一量的大小都是由相同种类的单位量来表示的；线长由线表示，角度的大小由角度值表示，属性总是由其个体来标志，一般性概念总是由其中某个具体对象来表示。通过这种方法，所有的矛盾均可避免，几何从来不允许被用来推知那些不存在或者不能显示的东西的原因。

普莱费尔（Playfair，1778，第318～319页）

以提供实例为特色的这种一致性证明的实际认证在19世纪甚至变得更明确，如果要说与以前的情形有区别的话。约翰·赫歇尔（John Herschel）的下述论断就表达了这种共同的态度：

如果将通过真理在具体事物中的应用作为真理的检验手段这一点放在一边不论，那么除了其自洽性之外，再也没有什么可以指导我们的认识了。但在公理性命题中，这一点等于什么都没检验……作为与真理同体的基本要素，它们（公理的）相互兼容性只能通过经验显示出来——通过其共存性作为在特定情况下产生的真理所观察到的事实显示出来。

赫歇尔（Herschel，1841，第220页，括号为本文作者所加）

弗雷格在反对创造主义者（creativist，如汉克尔和希尔伯特）学说时也强调了这一点。因此，他对汉克尔的一致性认定和通过观察确立存在的观点做了如下回应：

当然，严格来说，我们只能确立这样一种情形：一个概念，如果能先产生某个属于该概念范畴的事物，那么这个概念就是无矛盾的。相反的推论则是谬论，汉克尔的失误正在于此。

弗雷格（1884：1968，§95）

同样，在他与希尔伯特就他（希尔伯特）似乎相信存在无需提供可供验证的例证就可证明一个概念的一致性的可能性的通信中，弗雷格写道：

在谈到证明某些特性、要求（或者其他一些随你怎么称呼的品质）不能彼此矛盾时，我们的意思是什么呢？我认为它的意思只有一个：指向一个具有所有这些特性的对象，给出满足所有这些要求的事例。我们似乎不可能还有其他任何方式来证明没有矛盾。

1900年1月6日弗雷格给希尔伯特的信（Gabriel et al.，

eds，1980，第43页）

最后，受挫于希尔伯特不让步带来的沮丧，弗雷格要求希尔伯特解释清楚我们怎样才能够不通过确认可验证的实例就能够证明一个概念的一致性：

我相信我从你来信的某些地方可以推断出，我的论证没能说服你，这让我更渴望得到您的反驳。在我看来，你认为自己在证明无矛盾方面拥有一项原则，它与我在上封信里给出的原则有着本质的不同，而且，如果我记得没错的话，你的这项原则是唯一被你用于你的《几何基础》中的原则。如果在这个问题上你是正确的，那么它可能就非常重要了，虽然迄今为止我并不这么认为，但我觉得这一原则应该能够还原到我所制定的原则上来，因此它不可能比我的原则有更广泛的范围。这样做可能有助于厘清你在回复我上封信中提出的问题——我仍然希望得到一个回复——那就是你可以清楚地制定这样一个原则，并通过一个例子来讲清楚它的应用。

1900年9月16日弗雷格给希尔伯特的信（Gabriel et al.，

eds.，1980，第49页）

弗雷格在信中提到希尔伯特的《几何基础》，主要是指希尔伯特用解释实数的方法来证明他的几何系统的一致性（见Hilbert，1899，§9）。他正确地指出，这是希尔伯特在当时给出的唯一的对一致性的证明。差不多20年后，他的这一一致性证明理论概念才得到后人注意。[46]

因此，验证性实例的产生是确立概念一致性的唯一途径这一观念在17、18和19世纪里是一种普遍的信念。也正是这一观念使我们能够解释为什么莱布尼茨和其他人会将“没有A是B”的证明看作是对“所有A是

B”的证明的威胁。这个解释如下。如果“所有A是B”和“没有A是B”都是可证明的，并假设凡可证明的都是真实的，那么就不会存在A。但是，如果不存在A，那么我们实际上就没有办法确立所有关于A的断言的一致性（包括理论上的假设“所有A是B”）。这样做的唯一可行的方法是产生一个A的实例，并通过将它当作主体概念的断言来验证该实例具有归因于它的属性。最后，我们需要A的实例化，因为我们最终需要证明关于A的陈述的一致性，而作为一个实际必然性的问题，它只能通过产生一个A来解决。

对于数学是被创建的还是被发现的这个问题，这一推理的假想的结果是这样的：彻底的创造主义关于概念的观点是不可持续的。原因是，每个理论都必须保持一致，并且在最后的分析中，这种一致性只能通过识别见证的实例来显示。阿尔诺注意到，这种见证只能是被发现的，而不能被创建的。

> ……名义定义是任意的，而真实定义则不是这样。既然任何声音……能够表达任何思想，因此请允许我选择特定的声音来表达我自己的……想法。但这种情形与真实定义完全不同，因为获得各种概念的关系独立于人的意志。

阿尔诺（Arnauld，1964，第83页）

因此，名义定义和真实定义的适当角色的问题曾是关于发现与创建问题的传统讨论的核心。其核心要素被很好地传递到20世纪。我相信，了解这一点应当有助于我们深化和拓宽我们自己在这个问题上的观点。尤其是，它引导我们去看待问题背后的更多的东西，以及在它的解决方案中那些比哥德尔建议的论证更可能处于危险中的东西。

作为理论关切的真实定义

对真实定义的理论辩护基于我们如何获得概念的观点。这是一种广泛的经验论的或曰经验主义的观点，它将对经验的抽象看成是概念发展的根本途径。由此，它将概念归类为真实的或不按照对这些概念如何提出的问题是否存在一个合理的解释，是否存在一条从经验上升到抽象的路径来考虑。以这种方式提出的概念是合法的、真实的，而那些不是这样提出的概念则是不合法的、不真实的。

近代以来，这一观点得到了许多数学家和哲学家的提倡。例如，流行于19世纪初的约翰·莱斯利的几何教科书就教导说：“几何……是建立在外部观察的基础上的，但这种观察是如此熟悉和明显，以至于由此产生的主要概念给人一种直观的感觉，并被当作与生俱来的”（Leslie，1809，第2页）。他还说过，“几何学，像不关乎心灵活动的其他科学一样，最终取决于外部观察”（同上，第453页）。

在18、19世纪的思想家那里，“真实”有一个公认的代名词，叫“给定”，莱斯利将这个术语刻画为下列这句话（见上述引文的续篇）：

> 数量被认为是给定的，它可以是展现的，也可以是发现的。

莱斯利（Leslie，1821，第4页）

总的来说，几何学上的真实定义的理论辩护是基于两个观念。第一个观念是几何概念由抽象过程从对外部事物的观察得来；第二个观念是这种观察的内容在一个重要的意义上看是给定的而不是创建的。因此，真实概念是一种由给定内容开始的表达，而不是一种对被创建或制作的内容的表达。然后，它们经抽象过程从经验中提升出来，这个抽象过程只会从原始经验内容中减去而不是添加某些东西。

这种真实概念的观点有时会被19世纪的哲学家上升到关于概念的普遍性观点。一个突出的例子是叔本华，他对概念的性质以及它们的存在条件提出了以下令人难忘的解释：

> ……概念从直观领域派生出其内容，因此，思想世界的整个结构有赖于直观。因此我们必须能够从每一个概念——即便是间接地通过中介概念——回到该概念被从中抽象出来的直观……就是说，我们必须能够用在各种事例的关系中支持抽象的直观来支持它……
>
> 这些直观……担当得起我们所有思想真正的内容，无论何时，只要直观缺位，我们的头脑中便没了概念，只剩下单纯的文字。在这方面，我们的智慧就像发行票据的银行，如果它是健全的，它就必须有充足的现金来保障其安全，从而在要求兑现的情形下能够支付它所发行的所有票据。直观是现金，而概念则是票据。

叔本华（Schopenhauer，1859：1966，Book 1，第7章，第2卷）[47]

然而，这种通用型的观点局限于哲学家圈子里或仅流行于19世纪。事实上，直到最近数学家里也只有外尔（Hermann Weyl）形象地用存款票据来比喻真实概念与命题之间的差异，这是一方面，而另一方面，则是用纯粹的符号体系来描述二者间的关系。外尔用数学里的存在断言作为他的主要例子，他将这种断言比喻为纸币——就是说，钱本身没有任何价值，其唯一的真正价值是它背后所支撑的真实商品。

> 存在性命题——类似于“存在偶数”这样的命题——绝不是一个真正的事实判断意义上的断言。事物的存在状态是逻辑学家的一个空的发明。“2是一个偶数”：这是对一个给定判断的真实的事实表达。“存在偶数”则只是由这类判断得出的一个抽象判断，就像我将知识看成是一种有价值的财产一样，我将抽象判断看成是纸币，不论出现在什么地方，它都在某种程度上代表着财产。它唯一价值就在于它能够让我随时掌握着财产。如果没有像“2是一个偶数”这样的实际判断在背后支撑，这种纸币就不能变现，也就毫无价值可言。

外尔（Weyl，1921，第54页）

在上述所有事例中，我们看到有一个相同的基本理念在起作用，即我们称之为概念的那些表述只是这样一种存在：它们存在于这样一种可抽象性的关系当中，这种可抽象性是对那些给定的而非创建的经验内容的抽象。

这一判断为下述问题——为什么“没有A是B”的证明会被轻视，或“所有A是B”的证明会贬值——提供了一个比较明确的解释。如果“所有A是B”和“没有A是B”这两个陈述句的内容都是真实的，那么就不存在A。从而也不存在A可以从中被抽象的经验内容——就是说，不存在有关A的经验性实例。因此表达句“A”没有内容，也不会表达一个概念。其结果是，“所有A是B”不表达一个命题，因此也不具有真实判断的内容。因此关于“所有A是B”的证明不会产生所有A是B的知识，因此也就无法达到其预期的目的。

但是，尽管这种推理是明确的，它还是不能令人信服，这至少有两个原因。首先是它预先否决了存在非实例概念的可能性。这是令人担忧的，因为有可能存在那种既不是用实例来说明也不能用实例来说明的真

实概念（或至少是概念的陈述）。复合概念无论如何都属于这种情形（比如一张需要着5种颜色的地图对其着色来说就属这种情形）。是否存在那种能被正确地看作是原始的同时又不是可实例化的概念则是一个更难回答的问题。但据我所知，是否就不可能存在这样的内容也没有好的论证。

叔本华—外尔论证为什么不能令人信服的第二个原因，是它似乎无视这样一种可能性：能够具有经抽象导出概念的经验类型本身可能就预设了概念的可获得性。叔本华—外尔论证假设，我们可以有这样一类经验内容，它可以不经概念（不基于这种获得性的概念）就开始一个抽象过程。但是，要有说服力，就需要对内容及其抽象过程进行仔细分析。有了这种分析，我们才能够无须应用概念就把握一个至少有某类最基本内容的论点。此外，对于断言——存在一种由经验性内容经抽象而获得数学内容的合理途径——也必须给出一个说法。

只有到这种缝隙被填补之后，理论辩护才有可能被接受为一种对概念的存在性或可获得性的一般性解释。尽管作为数学概念的一般理论，这种理论辩护有缺陷，但它毕竟还是为数学中发现和创建之间的有价值的区分提供了一个基础。这个想法是，真实概念是一个可实例说明的概念，但非实例说明的概念的存在性不应被否决，其运用也不应被禁止，只是对它们的运用应有所控制，以便满足适当形式的一致性要求。这种一致性要求，具体来说就是，对于所涉及的概念可以证明满足不必一定要有实例来说明的条件。

讨论

上述两种对真实定义的传统辩护为数学中发现和创建之间的明确区分提供了诸多可能性。两种观点与哥德尔的现象学论证也有分歧。前者认为，发现和创造之间的区别缺少的不仅仅是对我们可能具有的对于强加于我们的真理的经验的一种现象学上的充分解释。它们没必要否定哥德尔的现象学推理背后的观点。虽然它们同样不必将这一点看成是经验在数学知识发展过程中的主要作用的指标。

两种辩护都持这样的观点：产生一个概念的实例本质上属于发现而非发明的问题。此外，它们都将发现看成是对那种确保一致性或可能性的实在的发现。另一方面，创建则不具有这种特性。因此，即使出于使

真实定义完成其预期的正当目的的考虑，也有必要将它看成是发现而不是创建。否则，数学家将可能面临着适当的正当要求的无穷尽的递归，而这正是数学家想要避免的情形。

我在上面提到，对真实定义的实际辩护似乎低估了一致性对所产生的概念可能采取的约束的形式的多样性。这是因为在希尔伯特的证明理论发展之前人们普遍认为，证明一个概念的一致性的唯一方法就是找到一个公认的实例。希尔伯特的证明理论则提供了一种语法类型的替代方法。

另一方面，哥德尔不完备性定理（特别是他的第二定理）暗示，希尔伯特的替代方法的可适用性是有限制的。这可能意味着，实际上，我们往往还必须依赖于产生一个实例或构建一种模型来证明一致性。结果，在有关证明一致性的各种实际可能性的问题上，弗雷格可能还是正确的。[48] 因此，在最后的分析中，我们很难确定这种实际辩护对发现的支持作用能够有多大。

结论

在本文结束前我想考虑最后一个问题：关于哥德尔提出的数学中给定性（givenness）或强迫性（forcedness）现象的运用问题。不过，这里我关心的不是我们能否可靠地检测到它，以及如果是的话，那么能在什么程度的信心和鉴别水平上做到这一点，而是什么东西能使对“给定性”或“强迫性”经验的恰当反应有可能或是应当被置于首位，假设我们事实上确实存在这种反应的话。我是被波尔查诺对数学中真实定义的角色和特征等问题的犀利质疑而带到这个问题上来的。[49]

众所周知，波尔查诺对中值定理的证明主要是他下述思想的产物：用来证明一个定理的方法应当具有使被证明的定理提升到一般化的水平。在他看来，关于数量的一般性定理的证明应该只求助于关于数量的一般性法则，而不应求助于仅适用于特定类型数量的法则。

这使他拒绝接受以前对中值定理的各种证明方法，因为他认为这些方法都是借助于具体的几何量来证明的。这也导致他拒绝接受通常的几何定理的证明，他相信其中的许多定理可以从有关数量的一般法则导出。这类定理的真正基础的一个更准确的图像可以由这样一些证明给

出：这些证明本身的前提就是数量的一般法则。

> ……在欧几里得几何中，没有任何空间对象是被作为真实对象接受的，除非它的结构首先借助于平面、圆和直线而被证实。这种限制显示出其经验来源十分清楚。黑板、圆规和直尺是……最初的绘图所需的最简单的工具。不过，就其自身考虑，直线、圆和平面都是复合对象，它们可能不能以任何方式作为一个假设被接受……例如，“在每两个点之间存在一个中点”这一命题就要比命题“每两个点之间可以作一条直线”简单得多。然而，欧几里得是用后者和其他一些命题来证明前者。我们证明了每一种被提出来的可能的概念性联系，这对于数学的理论论述是充分的……对象如何以及以什么方式类比到概念可以在实用数学的实在中产生。[50]
>
> 波尔查诺（Bolzano，1810：2004，§37）

在我看来，波尔查诺在这里做的部分工作就是指出了遵循强迫的引导的危险性。太多的信任有可能导致这样的事情：用并非基本的（但强迫程度较高的）命题来证明更为基本的（虽然强迫程度不高）的命题。

这就提出了一个与哥德尔的观点相关的问题。哥德尔认为，正是集合论公理迫使我们信其为真。同时他又将数学中的强迫性比喻成感官知觉。这使得我们认为，集合论公理对于更大范围的数学真理的重要性绝不亚于感性知觉对于更大范围的自然科学真理的重要性。

但如果这是正确的，那么根据数学和自然科学之间的类比，那些迫使我们信其为真的集合论命题就不大可能是数学的基本法则。也就是说，它们不应该被当作公理。相反，它们应该被看作需要由更深刻、更基本的法则来予以正确解释的现象，如果你愿意，你可以称它为理论数学。[51] 但我们很难设想这些更深层次的法则是什么。此外，集合论公理与感官判断之间的类比与哥德尔自己所说的关于集合论与数论之间关系的言论——即前者至少可以通过其简化后者的定理的证明的能力而得到部分的合理性证明——不是很一致。

总之，有关感性经验与哥德尔假设的由各种数学真理迫使我们信其为真所带来的经验之间的类比存在严重问题。如果说这种强迫是经验的显现，那么强加给我们的让我们信其为真的那些数学命题就应当被看成

是数学的经验部分。追随着哥德尔的这种类比，我们被引导到考虑数学中观察/理论分野的可能性，其方式类似于我们在自然科学中所接受的情形。波尔查诺在某种程度上相信这一点。具体来说，他认为各种数学命题之间存在客观的基础性差异。不仅如此，他还将揭示这种相对基础性的排序看成是数学家的主要职责。

由此，波尔查诺提出了有关给定性及其在数学认识论上的可能意义的重要问题。我已指出，这些问题对哥德尔试图引入一种经验到数学认识论的做法提出了挑战。它提醒我们应当对给定性或强迫性给予与基础性同等程度的重视。但就我目前所看到的，这些挑战还没法以同样的力度对数学中发现/创造问题的传统理解给出明确的意见——就是说，围绕所谓数学中“真实”定义的运用及其保护所展开的讨论也许提供了反对某种形式较为粗俗的主观性。

评迈克尔·德特勒夫森的“发现、发明和实在论”

约翰·波金霍尔

迈克尔·德特勒夫森仔细讨论了哥德尔提出的断言——我们有经验以一种无可商量的方式来面对数学对象，这种方式类似于我们在面对那种迫使我们将其看作物理世界的独立实在的物理对象时的情形。他非常正确地指出，这种断言，如果能够被证实的话，为许多数学家对所持信念的辩护提供了最佳基础，这种信念是，他们的推理过程是一个作出发现的过程，而非单纯的、如创制一种取悦性的智力拼图那样的发明过程。

就探索这种类比而言，我认为有必要充分考虑到我们与物理世界相互关系的微妙之处。实在论的捍卫者们（我自己也是其中之一，见Polkinghorne，1996，第2章）声称，现代物理学不是简单地诉诸这样一个事实：我们碰到的都是大的对象。例如，量子世界的性质就非常难以捉摸，以至于我们无法将其视为单纯的目标来处理。然而，物理学家们却对如电子或夸克这样的实体的实在性充满信心，不会将它们看成只是为了满足计算而想象出来的虚构的东西。我相信，对物理实在论的辩护最终取决于我们对这些实体的理解程度。我们以本体论的严肃性来看待电子和夸克，是因为它们的存在解释了大量更为直接可感知的现象。群

论在指明对称图案的结构（du Sautoy，2008）方面所显示出的力量似乎为赞成认真看待有限群的抽象实在性提供了一种类比。

对物理世界的实在性持赞成态度的另一种依据是物理世界经常表现出的令人惊讶的特性，这些特性总是与人们事先“合理的”期望相反。量子理论是这方面首屈一指的典范。这种对事先预期的阻挠雄辩地说明我们面对的是一种独立于我们的实在。我认为，19世纪对非欧几何的发现可看作数学上对此情形的一种类比。

德特勒夫森回顾了康德对直观（它以一种不容分辩的给定性作用于我们）与发明（他相信能够由我们的意愿创造或产生的东西）的区分。在我看来，这里与数学家证明了的、进入一种对深刻的数学理论的“完全形成的”意识状态的经验有一定联系。在我自己的那一章中，我提到过有关亨利·庞加莱的一段著名的轶事。庞加莱经过几个月对福克斯函数的持续不懈的攻坚后，在他正要启程去度假的当口，发现整个理论就像上帝送来的礼物一样突然呈现在他的脑海。数学思想的这种不容商量的特性似乎可以展现为下述事实：在考虑特定的公理化体系的数学家能够“看到”以哥德尔方式所呈现的真理，尽管它在该体系中形式上是不可证明的。

爱因斯坦曾经说过，基础物理学的基石是“自由发明”。他在说这番话时肯定不是要赞同那种将提出令人兴奋的概念归因于纯粹主观创造的后现代主义思想。爱因斯坦是以一种毫不妥协的客观性来思考这些问题的。同时，他也指出了创造性的想象力的飞跃（譬如）在他写下优美的广义相对论方程时所起的作用，尽管这些方程的结果还需要通过与现象的比较来验证。同样，数学中的那些伟大思想总是在显示“深刻的”性质时被掌握的——就是说，一个看似简单的公式被证明有着广泛的和意想不到的结果，正如一个成功的物理理论的预言所适用的范围之广一样，在该理论被提出之后人们才发现，其结果有着甚至出乎已知经验范围的广泛性。对此试想一下复数概念的提出所带来的惊人的丰富成果便可领悟。

实在论在为物理学和数学两方面的解释所做的辩护上必须是微妙而细腻的。在我看来，在这方面，这两个学科是一对“堂兄弟”。

第8章　数学与客观性

斯图尔特·夏皮罗

我想以一种试探性的和普遍的方式来探索在何种程度上数学可以认为是客观的这个问题。按照哲学上的典型做法，我们的目标之一是试图明确问题中各个术语的意义。我们当然知道数学是什么，或者至少当我们看到它时我们知道它是什么，当然那些模棱两可的个例除外。但什么是“客观性”呢？直观地说，所谓客观性就是独立于人的判断、规范以及生命形式等的一种物质属性。但“独立”这个概念又是什么意思呢？

认为数学不是客观的——数学真理总是与人类的认知、规范或其他东西联系在一起——的观点在哲学家中并不少见。伊曼努尔·康德将数学看成源自于“纯粹的直观”，是我们的感官从空间和时间上感知世界的一种形式。由此可见，数学直接与人的能力相联系，因此缺少客观性——至少在某种意义上说是如此。传统的直觉主义者L.E.J.布劳威尔和阿伦·海廷这样写道（Heyting，1931：1983，第52页）：

> 直觉主义数学家建议将数学研究作为人类理智的一种天然功能，是一种自由的、重要的思想活动。对他来说，数学是人的心灵的产物……我们不能将数学对象看成是一种独立于我们的思想的存在，即一种先验的存在……数学对象本质上依赖于人类的思想。

杰出的当代哲学家泰伦斯·霍根（Horgan，1994，第99页）采用希拉里·普特南所提倡的惯例，用小号大写字母词来表示“有关心灵独立、言谈独立（discourse—independent）的世界”。如果这个所谓心灵独立、言谈独立的世界是彻彻底底的物质世界，那么问题的实质就被回避了，但霍根和普特南都没有预先假设它这一点。

霍根声称（Horgan，1994，第100～101页）：

> 使得一个句子的正确的可断定性[52]能够依赖于这个世界[53]

> 可以有一系列的方法。在这个方法系列的一端是由……规范所支配的句子……（那些论述中的这种句子）仅当世界中的某个独特成分回答了每一个词项时才是真的……在方法系列的另一端是那种支配可断定性规范的句子……这类句子是被规范单独裁定为正确地可断定的句子，它们不依赖于事物如何与世界相联系。

随后他补充道，附带说明一下，“纯数学的句子”对于这第二种情形，即可断定的或“不依赖于事物如何与世界相联系”的情形，“是合理的候选者”。因此，霍根将纯数学缺乏客观性这一点看成是“合理的”而没做进一步讨论。

伽利略1623年出版的《试金者》[54]一书里有如下重要的段落：

> 哲学被写在宇宙这部永远呈现于我们眼前的大书上，但只有在学会并掌握书写它的语言和符号之后，我们才能读懂这本书。这本书是用数学语言写成的，符号是三角形、圆以及其他几何图形，没有它们的帮助，我们连一个字也读不懂；没有它们，我们就只能在黑暗的迷宫中徒劳地摸索。[55]

今天，我们是在此发表上述论断的17世纪头几十年更好的基础上来品味这段话所包含的深刻真理的。自然科学或社会科学几乎没有一个分支不是以大量的数学作为研究的先决条件。一个人如果没有相当雄厚的数学基础，就不可能跳过教科书的前几页。当然，这是指应用数学，而霍根评论的是纯数学。但是，某些科学的理论部分与纯数学之间的差异并不是十分明确，至少在我看来是如此。

像霍根这样否认纯数学的客观性的观点使得伽利略的观察成为一种奥秘——一种带有贬义的“奥秘”。如果像霍根认为的，数学不外乎受到人的可断定性的支配，那么为什么数学对于科学是如此重要？这个世界里一定有某种东西，就像这世界本身一样，是独立于人的关怀、判断等主观意愿的。这种主观意愿把数学置于我们试图理解的中心——即便我们的意愿是要理解宇宙。

伽利略用宇宙的语言说话。神学靠边站，当然我这只是个比喻。更准确地说，伽利略声称的是，为了全面或正确地理解宇宙，至少在科学上，我们必须求助于数学。那么接下来是不是就有一个数学断言

（assertions）本身是否客观的问题？这么说也可能没错：我们不可能在没理解一种语言的情形下就能够在任何复杂的深度上去理解宇宙。这样接下来是不是又有一个语言本身客观地、不依赖于人的兴趣、关注、判断和其他一些品质而直接与世界相联系的问题？不管怎么说，必定存在某种非语言世界的东西，这种东西使得语言成其为语言，并使这种语言成为与我们人类有效沟通的工具。同样，也必定存在某种非数学世界的东西，这种东西使得数学成为我们有效地——更是基础性地——理解其他东西的工具。

公正地说，海廷和霍根有一点肯定是正确的，那就是数学、语言和科学都是人类的活动，而这些活动的追求和结果要受到人类的关切和兴趣的影响。众所周知，不论是在数学还是在科学领域，理论和解释既带有非人类世界的性质，又与作为表述者和理解者的人的性质有关。正如约翰·伯吉斯和吉迪恩·罗森（Burgess and Rosen，1997，第240页）所说的那样，“我们关于生命、物质和数的理论很大程度上要受到我们的性格的影响，特别是我们的历史、社会和文化的影响。”当然，这并不是说，这个世界本身，或像霍根或普特南所说的，这个世界本身，受到“我们的性格”影响。伯吉斯—罗森的观察是，受到影响的是我们关于这个世界的理论。摆在我们面前的问题是：在何种程度上这些真理受到人类数学家的影响。

一些有竞争力的哲学传统认为，我们在将知识理论化时没有办法将“人”的因素和“世界”的因素完全分离开来。正如普罗塔哥拉（据说是）所说，“人是万物的尺度”。在某些唯心主义观点看来（更不用说那些后现代的观点了），世界本身就具有人的特征。世界是由我们的判断、观察行为等塑造的。因此，似乎根本就不存在霍根和普特南所说意义上的世界。一种不那么极端的立场是康德的学说：物自身是人类的探索无法接近的。我们是通过我们自己的范畴[56]、概念和直观来接近这个世界的。我们不可能超越自身而达到那个独立于我们的范畴、概念和直观的自在的世界（或世界）。

当代学术界广泛持有的观点是由奎因、普特南、戴维森和伯吉斯所倡导的观点。这种观点——如果用生硬的措辞来讲——就是，对于厘清一个成功的理论中哪些是“人”的因素使然，哪些是世界的贡献这个问题，根本就不存在上帝的观点，也不存在一种我们可以用之将我们关于世界的理论与这个世界本身进行比较的视角。

这种康德—奎因哲学取向可以说意味着根本不存在所谓的客观性，或者至少是不存在我们可以感知的客观性。如果这种观点是正确的，那么本文所提出的问题根本就没有答案。但不论这种观点有多少可取之处，我都会拒绝，尽管我对康德—奎因取向抱有同情。可能不存在诸如完全客观这样的问题——不论这种客观性指什么——但似乎仍然存在需要加以澄清和运用的既有趣又重要的客观性概念。例如在像“纯净的水分子含有2个氢原子和1个氧原子”这样的陈述与像“西兰花令人厌恶”“曼联队可恶”这样的陈述之间似乎有重要的区别。我们的问题是要搞清楚这种区分的哪一方包含数学。

克里斯宾·赖特的《真理与客观性》（C.Wright，1992）一书谈到过对客观性的一种解释，这种解释比我知道的其他解释更为全面，它提供了一种对基本概念的丰富而又详细的见解。赖特不是通过对实在的结构做形而上的研究来接近事物的性质，他不关心这个世界是否包含诸如道德属性或数之类的东西。他的注意力集中在分析各种言谈（discourse）的性质，研究这些言谈在我们的总体知识生活和社会生活中所发挥的作用。也就是说，赖特试图阐明，当我们试图谈论并理解我们发现自己所占据的这个世界时，对于我们——语言使用者的社会群体——来说，将一段言谈看成是客观的将意味着什么。

赖特认为，客观性不是一个意义明确的概念。客观性概念或客观性的基准（axis）形形色色，而一种既定的言谈只能用这种而非那种概念来表现。这些客观性的基准被冠以“认知约束”“认知律令”“尤西弗罗对比[57]”和“宇宙作用宽度”等名称。在以前的文章（Shapiro，2007）里我说过，除了在接近基础的地方可能存在某些令人不安的例外之外，数学很容易通过这全部四项检验。数学在认识论上是不受约束的：存在不可知的真理。伽利略的观察表明，数学具有极其宽广的宇宙学作用：它囊括了所有可能的解释，其中大部分都是对非数学问题的解释。在尤西弗罗对比方面，数学站在苏格拉底一边——它不依赖于响应——同时数学很容易满足认知律令。总之，如果说还有东西是客观的，那么数学就是客观的。

另一方面，可能的例外——基础性方面的问题——显得非常突出，因为这些问题涉及上面所提到的康德—奎因问题的核心。这里我想重新审视一下上述客观性的四项基准之一——认知律令，因为它证实了这一主题，其效果甚至超出了我在其他文章里的预期。

按照赖特的理解，只有当言谈受到认识上的限制时，认知律令才被认为是一项客观性的基准。因此，我在这里暂且稍事停顿，简单描述一下客观性的这项主要基准。

认知约束是迈克尔·达米特（Michael Dummett）的反实在论概念的一种表述。根据赖特对这项基准的表述（Wright，1992，第75页），言谈在下述情形下将受到认识上的限制：如果对言谈中的每个句子P有

P↔P是已知的

换句话说，如果一种言谈不包含不可知的真理，那么这种言谈就会表现出认知上的约束。

从“客观性”一词的本义似乎有理由得出这样的推论：如果认知约束对一个既定范围的言谈失效——如果存在这样的命题，在既定范围内其真理性不可能变得已知——那么这种言谈就只能有一种实在论的、客观的解释：

> 设想我们对一定言谈语境下的陈述的理解……通过给它们（指这些陈述——译注）指定这样的条件——潜在的超越证据的真理得到认可，即如果这个世界是按共同合作来运行的，那么任何此类陈述的真假都可以在我们的知识范围之外得到判定——而确立。这样……我们就不得不承认在下述二者之间存在区别：一种是不论在我们的言谈实践中采用什么标准都能使得这种陈述变为可以接受的状态，另一种是使陈述变成实际真理的状态。这种陈述的真理性的获得与我们采用什么标准无关……因此，达米特意义上的实在论是一种为下述观念奠定必要基础的方法：我们的思想渴望反映这样一种实在，其特征完全独立于我们和我们的认知操作。

赖特（Wright，1992，第4页）

换句话说，如果对于给定的言谈，认识论约束失效，那么这种言谈就是客观的，并且讨论到此结束。客观性的其他基准——认知律令、宇宙论作用和尤西弗罗对比——均与此无关，它们均不适用。另一方面，如果一个言谈受到认识上的限制——如果所有真理都是可知的——那么其他基准也将在客观性的重要方面显现。赖特就是这么论证的。但就目前的讨论而言，我们只需假设，在数学上，从“可知性”的某种意义上

说，所有真理都是可知的。因此我们转向对认知律令进行简要刻画。

假设一个既定范围的言谈被用来描述一个独立于心灵的世界的特征，以便进行直观的理解。假设有两个人在这方面不同意某件事情，那么我们有理由相信这两人中至少有一个歪曲了实在，这样，在他或她那里，对事情的评价一定在某个地方出错了。例如，假设两个人在争论一个给定的院子里是有7棵而不是8棵云杉树。假设这个言谈中在关于何为云杉树的认定上没有任何含糊不清之处，对院子的界定也没有模糊的地方，[58] 那么我们就有理由相信，争论的双方中至少有一个人犯了错误：他要不就是没看得足够仔细，譬如他的视力有问题；要不就是他不知道何为云杉树，他做了错误的推论；再不就是他数错了，或者是其他情形的错误。也就是说，存在争议这一事实表明，争议的双方必有一方有所谓认知缺陷（尽管要认定到底是哪一方有缺陷通常并不总是很容易弄清楚）。

相反，两个人可能在“这个宝宝是否可爱”或是“这个故事是否幽默”等问题上意见相左，这里就不涉及认知缺陷问题。两人中的一人可能在品位或幽默感方面很独特，或者也许根本就没有品位或幽默感，但这些都与他的认知能力没有关系。他可以和其他任何人一样感知事物和进行因果推理。

客观性的这项基准就取决于这种区别，取决于是否存在可以不受指责的分歧。赖特（Wright，1992，第92页）写道：

> 言谈表现为认知律令，当且仅当我们可以先验地判定，观点的区别只有根据“不同的输入”才能得到令人满意的解释，也就是说，争议双方是在不同的信息（因此负有无知或失误……的罪责）的基础上讨论，或是在“不合适的条件下”（导致注意力不集中或注意力旁置并由此造成推理错误，或看错数据，等等），或是在“出错状况下”（例如，对数据评估有偏见……或过于教条，或在其他范畴上出错……）讨论。

直观地说，在有关云杉树的言谈中，认知律令管用；在有关宝宝是否可爱和故事是否幽默的言谈中，认知律令失效。

那么在数学领域情况会怎么样呢？赖特认为，认知律令显然适用于

简单运算（第148页）。例如，假设两个人手工计算两个四位数的乘法，得出的结果不同。显然，两人中至少有一个出错。他可能忘了乘法口诀表（或者学习时就没有准确地记住它），再不就是运算时数位没对齐，或者注意力不集中，算错了。所有这些显然都属于认知缺陷。要全面评判赖特的这个结论，我们就不得不涉及维特根斯坦关于遵循规则的问题——这个问题似乎可以称为活动的客观性问题。但我们先把这些问题放在一边。

不管怎么说，数学里比简单运算复杂的问题多得多。认知律令是否对所有这些问题都适用呢？在专业数学研究中，一个严肃断言的认知标准是证明。因此，假设一位数学家，我们不妨称其为帕特，得出了一个他认为是对某个数学命题S的证明的东西，而另一个数学家克里斯对S心存疑虑，甚至在看了帕特提出的证明后仍持异见。那么帕特与克里斯之间的分歧就不必是证明结果上的分歧。他们对帕特所给出的证明是否够好——是否能建立起结论——这一点上看法迥异。帕特基于他的证明相信S成立，而克里斯不相信S成立，因为他根本就不认可帕特的证明的正确性。他可能认为S是假的，也可能他认为S根本就是不可知的。

在这种情况下，我们的问题是：我们是否能够先验地断定两位数学家里至少有一位表现出认知缺陷——假定分歧是真实的，至于到底是哪一方则是另一回事。

在专业数学的世界中，像这样的分歧经常发生。两个评审人可能对送审文章中的证明过程及其导出的结论持不同意见。这种分歧既不关涉评审人的能力，也不关涉论文作者的水平。但这种现象绝非数学所特有，任何足够复杂领域里的言谈都会有类似的“无可指责的”分歧存在。为了让认知律令有机会作为客观性的基准，并有助于获得有关纯数学状态的一些启发，我们必须将认知者理想化。[59]

在这里，问题的理想化可谓异曲同工。我们假定我们的对象在寿命、质料、记忆和注意力保持方面均无极限限制。在可计算性的数学理论和一般的数学逻辑里理想化指的就是这些性质。我们现在的讨论预示着有可能回到自身。为了评估认知律令在数学上是否成立，从而数学在这项基准上衡量是否客观，我们要谈及某些数学知识，使其理想化。我想大家都清楚，哲学做的是整体性研究。

对于数学的价值所在，在我看来，数学追求的更像是发现而不是发明，更像是获取真理而不是表达一种态度。作为一个有兴趣的局外人，在我看来还有一点需要指出，在大多数情形下，数学界呈现出一种明显的趋同倾向，或许是这种倾向在数学里表现得要比在其他领域更甚。从历史上看，至少在当下，有关某项论证是否正确的争论不会永远持续下去。除非争论双方失去了兴趣——那肯定是非认知性问题——否则实际的分歧似乎总是会得到令各方满意的解决。例如，到最后每个人都会同意，某些步骤是无效的，或是某项前提被取消，或是论证最终得到肯定。有确凿证据表明，在专业数学领域，认知律令成立，至少在适当的理想化情形下如此。

定义一个形式上的证明，就是用形式化语言给出一组句子，其中每一步给出的或是明确指出的公理，或是前提或假设，或是由前面的步骤通过原始的推断法则得到的其他什么东西（至少在作者看是如此），所有这些陈述是如此基本以至于做进一步分析已没有意义。在实际的数学研究中，我们并不总是明确一项给定的证明是否具有唯一的形式。在涉及如何将一段已发表的数学研究形式化方面，是否存在独立于判断及其类似性质陈述的客观事实呢？我承认，对此客观性的对手有一定的操控空间。从实际数学言谈——毕竟这是我们所关心的——到理想化数学家的完全形式化证明的变动可能不由完全客观的标准支配。但正如上文所述，在此情形下完全客观性不太可能做到。我们探索的是，在何种程度上数学在这项基准上看是客观的。

让我们回到我们想象的数学家帕特和克里斯的情形，当然现在的情形经过适当的理想化。假设帕特对数学命题S提出了一个完全形式化的证明Π，而克里斯拒绝接受这个证明，对其结论持异议。

假设帕特和克里斯对于出现在证明Π里的每一行公式都有一致的理解，那么这里的分歧就可能涉及其中一人身体上某个部分的认知缺陷，譬如他或她的眼神不好。这样，无论是帕特还是克里斯，都有可能在Π的某个公理、假设或前提下产生分歧，或是两人对帕特的原始推理规则的有效性达不成一致。

在数学上，表面上的对前提或公理理解上的分歧不是真正的分歧。两位数学家谈的都只是彼此的过去。帕特现在从事的是某种结构（或结构类型）研究，其部分特征由推导Π的前提刻画；而克里斯更青睐于另

一种结构。一个对毕达哥拉斯定理持有异议的数学家（因为他不假定平行公设成立）不会真的反对欧几里得几何。他们从事的是不同的理论、不同的题材的研究。

当然，数学家们并不总是这样想。据说，他们曾将有关几何的问题看成是关于（物理）空间结构或与知觉有关的直觉结构以及其他类似东西的问题。阿尔贝托·科法（Alberto Coffa，1986，第8页）描述过这种历史性转变：

> 在19世纪后半叶，借助于一种仍有待解释的过程，几何学界得出的结论：所有可能的几何都已明了……科学家群体以一种非临时性的方式就某一领域内一组互不协调理论的所有分支达成一致，这恐怕还是第一次出现……现在，哲学家们必须搞清楚……数学家对几何的态度的认识论意义……这项挑战对哲学家是一种艰难的考验，对此（悲观点说）他们都可能折戟……

我用当前的情景来重申这种观点。如果我说帕特和克里斯只是在论证到前提或公理方面意见相左，相信他们不会完全不同意。实际上，他们的话指向的是不同的东西。他们是在两种不同的结构框架下工作。这就解释了为什么数学理论在变得（至少不是现在）对科学无用后没有作为伪科学被丢弃的原因。迈克尔·雷斯尼克（Resnik，1997，第131页）称这种现象为“欧几里得救援”。

如果意见分歧涉及更为基础性的问题，那么事情可能就不会这么简明。假设“有争议的”证明的最终结果是一个实分析的命题，帕特的证明要用到集合论的原理，如连续统假设或大数假设，而克里斯不接受集合论原理。这样争论的焦点自然会集中到集合论这一背景理论上来。这里我们同样可以诉诸欧几里得救援，说帕特和克里斯是在不同的结构框架下研究，只是因为他们的集合论背景不同。譬如帕特用的是分析加集合论A，而克里斯偏爱用分析加集合论B。由于集合论的概念遍及数学的各个分支，加之集合论的基础性作用，因此这种区别并不像譬如欧氏几何与非欧几何之间区别那样一目了然（见Maddy，2007，第358～360页）。这个问题可以说是扑朔迷离，其程度已超出本文的讨论范围。我们不得不考虑数学是不是可以有不止一个的基础，并且，如果是这样的话，我们该如何搞清楚这些基础之间的关系。

不管怎么说，集合论背景下的争议就像是剩下的一种可能性，数学家所持态度的差异可以追溯到他们的逻辑。假设克里斯对推理法则持有异见，而这条法则在帕特看来已基本到不能做进一步分析的地步。譬如，假设帕特的证明要用到排中律，而克里斯拒绝排中律，因为他是个直觉主义者。这样就带来了一个逻辑的客观性的问题，这个问题可能需要（已经）由另一篇长篇论文（Shapiro，2000）来回答。个中细节已经超出了我们目前所关注的问题，但我认为，除了可能的认知约束之外，某些先决条件和逻辑已经经受住赖特的客观性检验。但我们应当从中得出什么结论目前尚不清楚。

这里一个潜在的麻烦是，赖特大部分的客观性基准都是按逻辑制定的，因此似乎是只有在我们确定了一种逻辑之后才可以运用各种基准。就是说，客观性的各种基准预先假定了一种逻辑（虽然是以哪一种逻辑为前提的问题尚未解决）。因此，我们甚至很难看出在何种程度上可以在赖特设定的框架下追问这种逻辑的客观性。

有人可能会援引欧几里得救援来处理逻辑，对它采取折中的态度。比方说，有论文认为传统的分析和直观的分析是两个不同的主题，它们之间的冲突并不比欧氏几何与非欧几何之间的冲突更激烈。如果我们继续沿着这条路线走下去，那么在一定意义上，所有的数学——一旦经过适当的理想化后——都可以还原为演算。这样问题就变成仅仅是对于不同的演绎系统可以得出什么样的结论的问题。在此之外数学没有任何内容。我们保留认知律令，但代价是数学家之间不存在任何有趣的、真正的纠纷。有可能起“纠纷”的双方用的不是同一种语言。我们似乎已经使自己陷入某种数学形式主义——至少是一旦经过适当的理想化，我们就可以运用客观性基准。当然，我们仍然将维特根斯坦的规则遵循问题放在了一边。

让我们简要地探讨一下另一种情形：古典数学家（克里斯）和直觉数学家（帕特）彼此间有真正的分歧。这时我们可以问，按照认知律令的基准，这个问题是否是客观的。这里要趟的浑水更多。

我们的问题是理想化的纠纷当事人——克里斯或帕特——是否会表现出认知缺陷。在赖特给出的上述认知律令特征中，缺陷表里有一项是“推理错误”。[60] 帕特当然会指责克里斯推理错误。克里斯援引排中律，但在帕特看来这是不对的。然而，赖特似乎没有考虑到这一点。他

将“推理错误”归因于“注意力不集中或注意力旁置”的结果，而不是一种涉及逻辑结果本身性质的深刻分歧。而我们在这里对人物已经做了理想化，克里斯和帕特都不会犯注意力不集中的错误。因此，如果认知律令在此成立，那么我们就不得不寻找另一些类型的认知缺陷来解释我们的理想化数学家之间的分歧。

可以说，逻辑选择是一种整体性行为，虽然这又是一个有争议且无法在此得到圆满解决的问题。纳尔逊·古德曼曾制定“宽的反应平衡”方案并被约翰·罗尔斯用于解释正义，雷斯尼克则提出了一种此方案的改编版：

> 我们从自身对逻辑正确性（或逻辑必然性）的直觉开始。这些直觉通常以一组测试案例的形式出现，譬如我们所接受或拒绝的论证，我们当作逻辑上必然的、不一致的或等价的陈述……，等等。然后我们试图建立一种合乎逻辑的理论，其陈述符合我们最初所考虑的判断。初步尝试就想在理论和“数据”之间形成精确的配合是不可能的……有时候……我们会形成我们自己的逻辑直觉以进行有力的或精准的系统考虑。总之，“理论”会使我们拒绝“数据”。此外，在决定我们必须放弃的东西时，我们不仅要考虑逻辑理论本身的优点……而且还应考虑该理论以及我们的直觉如何与我们的其他（包括哲学上的）信念和承诺相一致。当该理论不再拒绝我们所决心维护的任何例证，也不再支持我们所决心拒绝的任何例证时，该理论便与其最终的、考虑成熟的判断达到了……宽的反应平衡。

雷斯尼克（Resnik，1997，第159页）

这里存在另一种麻烦的循环论证。为了检查一个理论是否处于反应平衡，他必须进行逻辑推理。他必然会得出合乎其逻辑理论的逻辑结果，以便检验该结果是否与其直觉和其他“数据”相一致。我们不可能严格刻画反应平衡在逻辑选择上是否呈中性。我们只能针对一种既定逻辑来谈反应平衡。

不过，也许争论双方之一无法通过自身达到反应平衡。这肯定属于该争论者的一种认知缺陷。但是，我们可以肯定这种现象总是会发生吗？也就是说，我们能够排除认知上无可指责的分歧吗？

事实似乎并非如此，虽然我们很难看出人们是如何为此构建一项论

证的。有可能我们的两位理想化数学家——帕特和克里斯——依据各自的逻辑都是反应平衡的。如果是这样的话，我们又如何以中立观察者的身份来指责其中一方的认知缺陷呢？

在逻辑客观性的整个问题上也存在某种困扰。对于任何范围的言谈，严肃的争论都包含逻辑。无论什么范围，争论各方都有他自己的推理逻辑，并根据这种推理得出各自的结论。无论逻辑的应用范围有多广，有关逻辑的分歧或差异注定会导致分歧或差异出现在论证的任何地方。如果逻辑都失去了客观性，那么还有什么地方可以谈得上客观性？

我很抱歉，对于逻辑的客观性，以及某种程度上对数学的客观性，没能给出一个清晰的结论。由于逻辑在我们的理论化过程中起着核心作用，因此我们很难通过明快的处理将它区分开来。任何企图刻画客观性是如何被裁决的问题都必须以逻辑为先决条件。

如果我们还记得康德—奎因的论点，即我们没有办法将理论中的世界因素与人的认知因素截然分开来，那么这种情况就可以做得更合意。我认为，由此带来的一个结果是要求完全客观性是没有意义的。因此，某种程度上说，放弃特定范围的言谈的客观性并不会使客观性完全被消除。这是整体怪兽的性质。

赖特在他的书的后面（1992，第144页）对认知律令的制定增加了一些限定条件。这些条件看起来是要在为取得反应平衡而斗争中处理好整体裁决的问题。莱特认为，当且仅当下述条件成立时，言谈才行使认知律令：

> 这一点可以看作是先验的：除非像在有争议的陈述中，或在可接受性的标准中，或在个人的证据阈值出现变化时，模糊的结果是可以谅解的，否则在言谈内所形成的观点之间的差别可以说与某种能够被描述为一种认知缺陷的东西有关。

也就是说，赖特认为，那种取决于模糊性、可接受性标准等的无可指责的分歧不会损害认知律令。为什么会有这些例外？制定认知律令标准的最初动因究竟是怎样的，为什么认知律令没有以这种或那种方式提到模糊性、证据阈值等的问题？这是不是可以看作伊姆雷·拉卡托斯的所谓“怪物排除法[61]”的一个实例（Lakatos，1976）？就是说，我们一

旦发现理论的某些部分似乎不合适，就直接把它们排除出去。

尽管赖特没这样说，但我认为，认知律令的细致表述版本中列出的例外实例是符合康德—奎因论点的。按照赖特对模糊性的理解——也是我的观点（Sharpiro，2007a）——模糊词项是依赖于响应或判断的，至少在其边缘区域是这样。他写道，“很想这么说……一个包含（一种）模糊性的陈述表现为这样一个事实：某些情况下，认知上清晰的、得到充分了解并具有适当函项的主题在表述上可以相互各异且无可指责。”然而，在非常客观的研究领域，如自然科学领域，模糊词项是行得通的。仅仅因此并不能损害其客观性，除非我们以一种全有或全无的方式来处理这一问题。同样，在世界因素与人的认知因素混杂的理论中，“可接受性标准”，尤其是“个人证据阈值的变化”，更接近于“人”的一方而离“世界”较远是合理的。当然，保守的科学家，那些在提出断言时更为谨慎的人，较之那些从事纯理论研究的同事，不必有任何认知上的缺陷，反之亦然。因此，追溯到这种差异的分歧不必破坏认知律令。

有关数学及其逻辑的更基础的问题同样如此。指向整体性考虑的分歧最终可能会（譬如）部分地因为理论家对所发现的东西的趣味——讲求精致还是简单——而被裁定。某个数学家可能偏好构建性数学所产生的细微区别和较清晰的界限，而另一位数学家则可能追求统一，在某种程度上说，即追求经典数学的简单性和易于处理的元理论。也就是说，可能正是对于某些基础性问题，其探讨更接近于“人”的一方而非互联网上的“世界”一方。即使是这样，我们也并不能导出数学不是客观的，甚至认为数学不具有客观性范式——我们用以衡量其他言谈的标准之一——这样的结论。

评斯图尔特·夏皮罗的“数学与客观性”

吉迪恩 • 罗森

斯图尔特·夏皮罗间接地处理数学的客观性问题。他不是探求数学事实是否以某种方式依赖于人的思考，而是探求数学言谈展现的是否就是赖特所谓的认知律令。他在本章里预先假定这些观念都以一种简单的方式联系在一起：

如果言谈不能展示认知律令，那么言谈所涉及的事实就不是客观事

实。

在这篇评论中，我对这项原则提出疑问。

假设你我之间对瓦格纳的歌剧《尼伯龙根的指环》所展现的美看法不同。你说这是一部杰作，我说它很无聊。假设我们都认同歌剧及其背景音乐非常优美，观看演出时我们也都没喝醉或是注意力不集中，我们的判断作为对外界的反应是稳定的，那么，即便我们双方都牢牢把握基本事实，双方都没犯推理上的错误，我们仍会有不同的见解。迫于要解释这种分歧，我们可能会将问题简化认为我们之间的分歧仅仅是因为我们对这个问题存在着彼此不相容但各自协同一致的审美感受。

粗略地讲，如果每一项分歧要不就是不同的输入——即争议双方可获取的信息有差别——的结果，要不就是某种认知障碍或推理出错的结果，那么认知律令就会在某个地方显现。上面这个例子表明审美言谈（aesthetic discourse）没能展现认知律令，因此鉴于上述前提，这个审美事实就不是客观事实。夏皮罗认为（有一些限定条件），数学言谈以完胜的姿态通过了检验，因此就目前的这一标准而言，我们没有任何理由怀疑数学的客观性。

这是正确的吗？近代数学已经是一个从公理出发通过规则来证明定理的体系，因此任何不是由简单错误造成的数学分歧总可以追溯到要不就是公理方面的分歧，要不就是逻辑规则方面的分歧。让我们暂时将逻辑分歧放在一边。逻辑分歧是一个重要问题，但正如夏皮罗所指出的，这个问题很难在赖特的框架内讨论。如果我们聚焦于公理方面的分歧，那么首先要强调的（正如夏皮罗做的那样）就是，虽然这种分歧偶尔在数学里出现（例如，关于几何的平行公设的争论），但近代数学已经有一套解决这个问题的标准方法。这个想法是把有争议的公理当作特殊数学结构的定义条款。譬如几何学家曾对平行公设的绝对真理或虚假的性质有不同的意见，而近代数学家则会说："有些空间是欧氏的，有些则不是，这条公理之所以在每个欧氏空间下成立，是因为满足该公理是构成欧氏空间的条件的一部分。但它在其他空间下未必成立，这方面例子很容易给出。"对于这个解释，探求这条公理是真还是假是没有意义的，因此对其真理性的认识是否达成一致也没有意义。

数学上的表观冲突通常可通过这种方式来解决，但情形并不总是这

样。数学的一个核心问题是确立满足各种条件的数学对象的存在性。现在，存在性的证明总是需要至少一条存在性公理。如果某人检验实数域下的数学，会发现我们总是能够简单合法地断言存在自然数和某些自然数的集合。这些存在性断言不单纯是条件或假设。当数学家通过构造（譬如说）R^3 上的一组公理（模型）证明了双曲几何的自洽性，他的定理的内容是："如果存在数，那么一定存在一个公理模型。"他的证明彻底地确立了模型的存在性，这意味着它必然包含至少有一条存在性公理断言。

当然，对于自然数以及由自然数构造的某些数集的存在性，数学界并没有真正的分歧。但在应用认知律令检验时，我们并没限定只专注于有关实数的分歧。即使对于《尼伯龙根的指环》的华美没有实际的分歧（感谢统一的音乐教育），仅仅是上述可能的分歧也足以表明，审美言谈通不过客观性的认知律令检验。同样道理，对于标准数学的存在性要求，仅仅是可能的分歧就足以确立至少在数学的某一部分是存在非客观性问题的，只要这种分歧不会被追溯到"不同的输入"或"推理错误"。

这种分歧显然是可能的。由传统方式培养的数学家发现，基本算术的存在性断言是一望便知的，因此无须证明就可接受。但我们知道，有可能某个人就是认为这些公理不是显然的。毕竟，就有哲学家基于下述两点明确拒绝将其视为当然：（1）它们没有内在的合理性；（2）有利于这一断言的每一项支持性论据均缺乏说服力（*Field*，1980；另见本书玛丽·伦的文章）。例如，这些哲学家通常会指出，数是某种不可见的、非物质的实体，这种东西的存在性不是显然的。

我想可以将这一点想象成一种真正的感性（*sensibility*）冲突。有些人认为存在性公理是显然的，因此对其持肯定态度，另一些人认为存在性公理并不完全是显然的，故持保留意见。这算不算一个"不同输入"的问题？当然，公理的情形也不尽相同。但如果我们在运用认知律令标准时把它当作不同输入的问题，那么我们不得不说我们在瓦格纳的歌剧问题上的分歧也是一种不同输入的问题，这将使标准变得无足轻重。推理中的错误或其他一些认知障碍算不算分歧？也许算，但如果把它归结为公理合理性的内在的分歧，我们很难看出为什么应该是这样。

在标准的数学的存在性假设问题上存在各种可能的分歧这一点表明，普通数学可能不会表现出认知律令。由此我们是不是就能得出数学

毕竟不是客观的这个结论呢？不能。我们能得出的结论是认知律令是一项有缺陷的客观性标准。客观性在这种意义上是一种形而上学的概念。所谓某个事实是客观的是指它不依赖于任何有趣的思想或语言方式（Rosen，1994）。从这一事实——在“数学合理性”的意义上说数学分歧的变化是可追溯的——我们得不出有关这个问题的形而上学方面的任何结论。如果我们发现，有关上帝的存在性的分歧有时可追溯到神学“感性”上的差异，那么我们是否能由此得出结论：上帝的存在（或不存在）在某种程度上依赖于我们的心灵？我们可以得出的结论只能是，我们关于上帝存在的判断并不是由证据严格强加给我们的。但是，这个结论的要点是这些判断的认知状态，而不是与它们有关的事实的形而上的性质。

答复吉迪恩·罗森

斯图尔特·夏皮罗

我倾向于同意吉迪恩·罗森对我的“数学与客观性”一文的评论意见。特别是，我同意他的这一观点：克里斯宾·赖特的认知律令的概念本身不提供客观性的必要和充分条件。我的这篇文章（以及两篇其他文章）的目的是检验数学与赖特有关客观性的不同基准之间是否相互抵触。不过我认为，关于认知律令作为至少是一种可撤销的标准有其正确的地方，这是一个阐明其范围和限制的问题。我同意赖特的这一观点：客观性不是一个意义明确的概念。存在各种方面的客观性，不是所有这些客观性都能够彼此排在一起。

在另一篇题为“客观性、解释和认知不足”的文章里（见待出版的克里斯宾·赖特纪念文集），我提出了这样一个思想实验：两位科学家彼此意见相左，但每个人在有关证据的总体平衡方面都处于反应性平衡状态。两个人都不会轻易出错。然而我们不愿意就此得出这样的结论：科学一般来说不是客观的。这篇文章的评审人指出，赖特的标准本身并不能对因主观性问题（如罗森关于瓦格纳的歌剧《尼伯龙根的指环》的例子）的非客观性质所引起的认知律令缺陷与因证据贫乏而造成的缺陷加以区分，这在需要对证据进行整体评估的问题上尤为如此。也许对罗森关于数学存在性的例子和虚构主义的一般性哲学主题，都可以做同样的理解。文章评审人是从其他不成功的尝试划分认知意义的历史的角度来

看待这个问题的。

在他的评论的最后，罗森暗示，认知律令似乎没站对位置：客观性是一个纯粹的形而上学的问题，而认知律令是一项具有广泛意义的认知标准。我不同意这一观点，但这里不是讨论这个一般性问题的地方。

第9章　数学对象的实在性

吉迪恩·罗森

如果明了真相，就明白不存在数这种东西。这不是说在15和20之间不存在至少两个素数。

保罗·贝纳塞拉夫，“数不可能是什么”（Benacerraf，1965）

问题

保罗·贝纳塞拉夫的著名论文的最后这句话是一桩公案：看似废话，但点出了——或者说似乎点出了——一个深刻的真理。在15和20之间存在[62]两个素数，这是一条基本的算术定理。因此任何承认算术基本定理的人都必然同意在15和20之间存在两个素数，因此至少存在两个数，因此存在多个数。然而数是真实存在的想法——真实世界包含数学对象就像包含枪和兔子一样自然的想法——听起来会让人觉得荒谬或困惑。因此我们发现哲学家总在使劲地阐明以下立场：

> *当然也存在数（以及函数、集和其他各种数学对象）。*
>
> *但这只是在数学里，这一点我们没有争议。而在另一种意义上——形而上学的意义上，并不存在数。数不是真实的存在。数不是任何一种具体的东西。*[63]

显而易见，这种言论像其立场一样相当令人费解。假设灭虫专家告诉你，你家阁楼上有松鼠，然后他又继续补充道，当然，在严格的和哲学的意义上说，不存在松鼠。或者假设某个天体物理学家报告说，在某个星系的中心有三个黑洞，然后又说：“顺便提一句，黑洞不是真的，它们不是具体的东西。”如果你像我一样，你会发现这些言论不可理解。然而在数学哲学上有一种普遍的意识：这种阐述方式，虽然可能不是像人们乐于接受的那样明晰，必须在某种意义上得到理解。

本文试图阐明什么是有条件的数学实在论。这种实在论之所以是实在论的一种形式，是因为它坚持认为数学对象是存在的。而之所以称之为有条件的实在论，是因为它认为这些数学对象毕竟在形而上学看来是属于“第二等”的。困难在于这种条件指的是什么。

我一度曾认为，以这样一种方式来产生一种值得讨论的有条件的实在论是无法做到的（Rosen and Burgess，2005；Rosen，2006），现在我认为这种悲观情绪是不成熟的表现。本文的目的是对有条件的实在论者可能心中所想的那些东西给出一种解释。它是这样一种解释，按照这种解释，从某种形而上学的观点看，数学对象是与某种范式——日常经验的对象，也许或是物理科学的对象——作不恰当的比较。我并不赞同下面将要讨论的有条件实在论的观点，但我相信这种观点值得我们注意，其目的只是要把这种观点摆上桌面。

关于数学对象的简约实在论的事例

在我们开始之前，综述一下实在论本身的事例（Burgess，1983；Burgess and Rosen，1997；Rosen and Burgess，2005）也许有助于我们对后文的理解。

所谓有关数学对象的简约实在论，我指的是断言数学对象存在的实在论。这里的“数学对象”可谓包罗万象，包括数、函数、集合、群、空间、模型、向量、范畴、方程组、形式语言和数学涉及的其他各方面对象。我们说数学对象存在，这句话的意思是至少存在一个这种类型的对象，或者说，这样的东西是存在的。

在这种语境下我一定要用“exist（存在）”这个词吗？我不认为是这样。与“存在”有关的习语——谓词“exists”，以及像“there are ……”，“there exist ……”，“at least one ……”等定量表达式——都是数学的日常语言的一部分。这些习语在语言上都是等价的，它们的含义很清楚。这就是为什么在标准的数学形式化表达中它们都可以用一个符号“∃” 来表示的原因。如果你正在阅读本文，说明你已经明白这种语言，我建议你就按照这种理解来把握全文。因此我想补充一点，当我说简约实在论指的是“数学对象存在”这样一种观点时，我是在通常的数学意义——即一个高中学生或一位职业数学家在谈到某个方程有两个解时心里所想的那种意义上来使用“exist”这个词的。它不是一个定义，但它足以

满足本文的需要。[64]

按照这样的理解，简约实在论就不是一种深奥的形而上学的断言，而是一种数学上的断言——一个有关数学最基本部分的平凡结果。正如前文指出的，它是一条表明在15和20之间有两个素数的算术定理。这条定理要求必须至少存在两个数，因此至少存在两个数学对象。因此任何接受算术基本知识的人除了接受简约实在论别无选择。

我们应当接受算术基本知识吗？我们有什么理由要相信在15和20之间有两个素数？在我看来，初等数学的断言与其他方面的常识性断言——例如断言我们生活在一个真实事物（即使我们没注意到它们，它们仍存在在那里）的世界里，再譬如断言生活在与我们相同的环境里的其他人类也具有有意识的心理活动——有同样的地位。当然，常识容易出错，但是，如果一个哲学家（或科学家或其他人）希望对常识性断言提出疑问，那么他必须给出怀疑的理由。而就算术基本知识而言，这样的理由根本不存在。算术在数学基础上显然是无懈可击的。[65] 如果我们将接受或拒绝一个理论的普通的科学标准作为权威，那么每一门成熟的科学都将运用数学视为当然这一事实就足以表明，广义上的科学活动从来没有对算术提出过怀疑。如果说在这一领域还存在任何怀疑的理由，那么这些理由必定出自独特的哲学理由。我不想对哲学家们在这种语境下所发展的论据进行点评，但我会（很简单地）谈谈为什么我觉得它们缺乏说服力。现有的论据分为两类。一些哲学家认为，我们应该拒绝标准数学，因为存在数和数类是件很奇怪的事情。这确实是真的。如果存在26这样的东西，那么它既明显不同于我们日常经验的对象（如桌子等），也有别于物理学所揭示的那些不为常人所熟悉的对象（如夸克等）。那它是什么呢？要想以这个理由来倡导拒绝算术，那么宏大的形而上学体系（有时也被称为物理主义，按照这个学说，一切绝对存在的事物都应像桌子或夸克那样真实）要比像15和20之间存在两个素数这样的算术断言更让人信赖。但是如果问题是在宏大的形而上学体系与算术基本知识之间进行取舍，那么在我看来，很明显，形而上学应当作出让步。

另一类哲学家拒绝基本算术是因为他们认为，如果存在数和其他数学对象，那么我们就没有办法知道关于它们的任何东西。这种断言通常得到关于知识的一般哲学理论的支持。根据这一理论，获取知识需要通过知识获取者与其研究对象之间的某种形式的互动才能实现

（Benacerraf，1973）。这些理论最初被发展出来用于解释经验知识，从这个目的来看，它们也许是有用的。但如果哲学家坚持认为它们对于完全一般性的知识仍能够成立，那么他就将面临一种严峻的困难境地。这些约束性理论通常会要求，确定数学观点的通常方式——计算、证明、非形式数学论证——不可能成为知识来源（因为它们不涉及研究者与数之间的因果互动），因此，那些以通常方式相信235+657=892的人其实并不真正理解为什么235+657=892。但这样问题就来了：为什么这个例子不能看作是对这些哲学家的理论的一个简单的反例？在其他领域，当一种哲学理论与外界公认的事实不相容时，通常的反应是重新考虑理论或重新限定其适用范围。知识哲学理论必须能够容纳数学知识这一显然的事实。如果理论与这一事实有冲突，那只能说这种理论没用到家了。

这只是对复杂的辩证逻辑的一种粗略的勾勒，但其中的主要策略应该是明确的。不论是以数学的相关准则来衡量，还是以科学和常识的标准来衡量，数学中的核心断言都是无可非议的。因此，对这些断言的任何哲学挑战都只是某种特定的怀疑论性质的挑战，这种挑战依赖于明确地将哲学原理应用于其本身并非纯哲学问题的问题上。这种怀疑论挑战出了名的软弱。当思辨哲学与公认的科学知识或常识有矛盾时，正常的反应——我相信这也是合理的反应——是怀疑哲学家是不是错了。这不是一项颠扑不破的原则，但却是一条很好的经验法则。如果在目前的这种语境下我们遵循这一法则，那么我们别无选择，只能得出结论：因为在15和20之间存在两个素数，因此，数学对象存在。

有条件的实在论：一个例子

这样看，简约实在论不只是一种哲学上的断言，而是公认的数学的一个（微不足道的）部分。但对许多哲学家来说，这种观点让人形成一种令人反感的图像——数学像是一门动物学，一门以描述特定类别事物的奇妙行为为目的的科学。二者的主要区别在于，在数学里，我们感兴趣的对象是数和那些完全不可见事物的无限性。[66] 正是这种思想，包括其他一些思想，使得哲学家沿着我们称之为有条件实在论的方向探索，即形成数不是在狮子和老虎这类东西的意义上为真的东西。我们的任务是要给这种暗思想一种意义。

我们从一些似乎鼓励这种空话的观点下的例子开始。这些例子全是

各种版本的算术还原论。一般来说，还原论在这方面的论点是站得住脚的：某种程度上，算术事实就是最基本的事实，没有其他更基本的事实作为其基础。现在，与此有关的一些还原论者的建议已不具有有趣的形而上的意义。如果理论家用纯集合的概念来标识具体的数——譬如称0是空集，接下来，数n是一个集合，它有唯一一个元素是n——那么算术的每一个事实都可以还原为集合论的事实。然而，这种还原论本身没有要抨击数的实在性的倾向。毕竟，它与下述观点是一致的：集合的概念与任何类似于实际事物的可能的东西一样坚挺，由于这种观点将数等同于集合，因此它与有关数的无条件的、十分牢固的实在论是一致的。

当我们考虑那些旨在将数学某些部分的真理还原为那种能用某种更基本的词汇（即那种不会将数学对象等同于更基本的理论所认同的对象的词汇）来表述的真理的建议时，我们便得到一种更有趣的还原论形式。作为证明，我们来考虑算术哲学里的形式主义。形式主义哲学的核心思想是，算术最终只与事物的语言域相联系。相反，只要算术有适当的题材，它便是算术语言本身及其句式之间的某种形式上的关系。 [67] 这里给出一个这种观点的简单版本。令PA是初等算术的通常形式：一阶皮亚诺算术。 [68] 设PA_{ω} 是这一理论的升级版（补足了推理的无限性法则——奥米伽法则（ω—rule），即允许推理从无限序列前提A（0），A（1），…，A（n）…推得普适量化的结论：对于所有的数x，A（x），PA_{ω} 显然是一个完好的理论。公理是真实的，推理法则维护真理。更重要的是，PA_{ω} 在下述意义上也是一个完备的理论：在算术语言中，每个句子A具有这样的性质：A或A的否定句总有一个在PA_{ω} 中是可证明的。 [69] 这意味着，任何接受标准算术的人都应接受以下的等价陈述：

对于算术语言的任何句子A， A为真当且仅当A在PA_{ω} 中是可证明的。

就其本身而言，这种等价性是一个没有什么特别的形而上学意义的数学事实。即使是最顽固守旧的柏拉图主义者也会接受它。现在我们来考虑具有形式主义特征的断言：

对于任意真的算术句子A：

A为真，因为A在PA_ω 中是可证明的；或者

使A为真的东西在PA_ω 中是可证明的事实；或者

A为真是基于这样一个事实：A在PA_ω 中是可证明的；

A为真包含在A在PA_ω 中是可证明的这一事实中。

这里的楷体语词不是数学的正式词汇的一部分。数学使我们确信，有关自然数的断言等价于某个形式体系下某个句子的形式可证明性的断言。但形式主义的独特的断言是：算术事实是以某种方式建立在这些证明理论的事实基础之上的，因此相应来说，证明理论的事实更为基本。从数学的角度来看，这种断言很不专业。它是一种关于算术形而上学的独特的哲学断言。

我们注意到，即使还原论性质的形式主义者将语言学上的某些成分（句子和形式系统）看作比数更基本，他也不能否认数的存在。由于算术的存在定理（例如，在15和20之间存在两个素数）在PA_ω 中都是可证明的，因此形式主义者必须承认这些定理都是真的。但正如我们已看到的，如果这个特定的定理是真的，那么就存在两个素数。因此形式主义者必须接受（就像人们喜欢强调的那样）存在数，数存在。就我们的目的而言最重要的是他接下去要说的东西，即全盘接受数的存在等于某些句子能够以某种形式计算方式从某些其他句子导出。正是在这一点上，他开始变得不像有关数的十分牢固的实在论者。

对此人们试图解释如下。如果数是第一位的，那么我们关于数的断言就可能是真的，这些断言为真的部分原因是数有它自己的存在方式。看看动物学就知道这是怎么回事。企鹅是实实在在的东西，关于它们的真实断言之所以为真，至少部分原因在鸟本身。相比之下，在如形式主义所理解的算术中，关于数的陈述之所以为真（当此时），是因为这种陈述能够以某种形式规则从某些公理中推导出来。数本身在奠定算术真理的过程中没起任何作用。事实上，人们通常很自然地假设这种因果关系是倒过来的。如果我们要求形式主义者告诉我们为什么关于素数无穷的欧几里得定理是真的，他会说，这之所以是真的，是因为它是一条PA_ω 定理。如果我们接着问他为什么存在无穷多个素数——也就是说，如果我们问的是有关数本身的问题，而不是关于某个句子是否为真的问

题——他可能会说：存在无穷多个素数，是因为“存在无穷多个素数”[70] 这句话是真的。按照这种解释（“有点超出上面所表述的形式主义的明确观点）[71]，不仅数学对象在使我们的数学理论为真的过程中没有发挥作用，而且对象本身的存在也仅仅是因为我们出于其他理由给出的关于它们的断言为真。只要听到这一点，便明显可知形式主义就相当于一种关于数的有条件的实在论。

另一个例子：模态结构

作为另一个例子，我们来考虑关于算术基础的一种结构主义版本。结构主义始于一种重要的观点：只要我们证明了关于自然数的一项定理，我们事实上就已证明了关于同构于数的对象的任一集合的更一般的定理。自然数的标准顺序为

0，1，2，……

我们用例子来给出一种独特的模式：一个离散的没有最后元素的线性序，其中每一项只有有限多个前项。如果罗马皇帝的继承能够永远延续下去，则序列

奥古斯丁、提比略、盖尤、……

将是这种模式的另一个实例。我们称像这样的任何有序集为ω序列。一般情况下，有序集是由集合X和这个集合上的关系≺组成对（X，≺）。在上述假设的例子中，各朝代的罗马皇帝之间有时间关系x的统治在y的统治之前，他们共同构成一个ω序列，就如同具有x小于y关系的自然数构成一个ω序列一样。

为了说明与算术有关的相关事实，我们需要一种更普遍的观点。众所周知，如果给定合适的逻辑背景，那么每一项能够用现代数论的精准专业语言陈述的算术断言都可以用仅有原始符号N（表示自然数）和＜（标准的“小于”关系符号）的精练语言来表示。如果A是一个普通的算术语句（例如，在15和20之间有两个素数），则我们可以将它翻译成这种精练的语言A[N，＜]。

相关的定理如下：

对于任何用算术语言陈述的断言A， A[N，＜]当且仅当对于任何ω序列（X，≺]，A[X，≺]。

这里A[X，≺]是在A[N，＜]中用变量X和集合上的关系≺ 替换专门的算术词汇N和＜的结果。因此，如果A[N，＜]是用算术语言给出的一个描述自然数展现某种算术功能的句子，而等价关系的右边是说，任何ω序列——即使它的元素是罗马皇帝——展现出相应的纯结构特征。因此，该定理要求：每一个关于数的算术断言等价于一个关于ω序列的完全一般性的断言。

现在，在这一点上，我们必须注意这个小定理的假设。ω序列是一个无限集合，因此，如果不存在无限集合，就不存在ω序列。但如果不存在ω序列，那么等价关系右边的每一个实例就是平凡真实的。这意味着，这个等价关系仅当实际存在无限集合时才成立，人们可能希望避免做这种假设。要做到这一点的方法之一是考虑一个有点不同的等价关系。即使事实上不存在无限集合，但原本是有可能存在的。罗马帝国没有永远延续下去，但原本是可以延续的。考虑到这一点，我们可以考虑以下等价关系：

A[N，＜]当且仅当作为必然性问题，对于任何ω序列（X，≺]，A[X，≺]。

即使真实世界里不存在无限集合，但这样的断言——任何可能的ω序列都具有某种结构特征A[X，≺]——则是不平凡的（在假设有可能存在ω序列的基础上）。这一定理肯定了这种模态断言等价于我们开始时说的普通的数学断言。（一个模式断言是指描述什么是可能的或是必然的断言。）

到现在为止，这仅仅是一个没有争议的（如果有些陌生的话）数学。任何接受标准算术的人都应该接受这种等价性。结构主义者特有的哲学断言是，算术真理可还原到或者说是基于关于所有可能的ω序列的一般性断言。[72] 回到我们的“在15和20之间存在两个素数”断言的例子。结构主义者接受这个断言，是因为他认为普通算术满足其上述立场。这意味着他（如同形式主义一样）不能否认数的存在。他的独特断言是：关于存在数的这个事实是基于如下事实：如果存在任何一种ω序列，就一定存在某种复杂的结构性质——粗略地说，对于此例就是在第

16个和第21个元素之间存在两个“素元素”。[73] 这里最关键的一点是：从表面上看，这后一个事实不是关于一种具体类型对象的事实。具体来说，它既不是关于数的事实，也不是关于罗马皇帝的事实。事实上，它让人感觉到这个事实不关乎任何东西。毕竟，形如“如果存在某种无穷集合，那么它将显现这样和这种性质”这样的条件模态断言不确认真实世界中任何事物的存在性。根据结构主义学说，算术真理正是由这种模态事实赋予的。这就不难明白为什么那些接受这一观点的人可能会不由自主地说有一种数不是真实的存在的感觉。关于企鹅的真理是由鸟类及其行为等事实而变得为真的。相反，有关数的真理并不是真正由数赋予的，而是由数本身并不显现的一般性条件的事实赋予的。

制定建议的框架

例子可以倍增，但其模式应当是明确的。对于给定种类的数学对象（例如，自然数），当我们凭借更基本的语词所描述的某种真理（问题中的对象不在其中显现）得到关于这些对象的每一条真理时，我们说这类数学对象是可还原的。在我们的例子中，还原性事实——即关于PA_{ω}中可证明性的事实，或关于每一种可能的ω序列——并不明确涉及（或量化）数，而借助于这种还原性事实，我们确立了有关数的每一个事实。如果这种形式的观点是正确的，那么当在某种意义上说数是完全真实的存在——存在这种东西——时，它们在下述意义上倒可能是“不真实的”：当我们检查最终基于算术真理的事实时，我们发现不存在任何种类的数。这提出了一个口号：真实的东西还原后不会消失。正如普特南指出的（Putnam，1967），拒绝“数学对象图像”的一种方法就是认为数学对象仅在这个意义上是不真实的。

讨论到这里已经将一种神秘性变换成了另一种。我们开始时想搞清楚的是，数和其他数学对象是不是真实的存在物。这个最初的设想是想用其他概念来解释这个概念——一种真理可以还原为另一种真理的思想。然而这个概念非常有问题。当然，在我们（譬如）说将有关意识的精神生活的真理还原为有关大脑和身体的生理过程的真理时，这句话到底是什么意思，可谓见仁见智，很混乱。还原的概念有许多，因此解释我们的建议有许多种方法。与调查选项的方法不同，我喜欢简单地勾画一个概念，我认为这种方法特别适合于我们的目的。

按我的理解，还原实质上是各种事实（而不是各种句子或陈述）之

间的一种关系，尤其是那些句子旨在描述的诸项事实或状态之间的一种关系。因此，还原是一种形而上的关系，而不是一种语义关系。说一种事实还原到另一种事实不是达成一个关于语词意义的断言，而是达成对事实本身的断言，这些事实的确立通常完全独立于我们的描述和思考它们的能力。

就目前而言，我们应将事实看成是由对象、性质、关系和其他各种属性所建立起来的复杂实体，这里说的属性与用词构建的句子有大致相同的意义。例如，2+3=5就是一个复杂的事实，它可以表示为：

[=（+（2，3），5）]

在这里，数字2和3以及等号和加法运算都是这个事实的组分，正如“2”和“3”是句子“2+3=5”的组分一样。[74]（在下文中，方括号中的句子命名为一个事实，其结构对应于封闭句子的结构。）

我的主要的实质性假设是：事实由基础的基本关系确立。这种关系没有标准的英语单词可加以描述，也没有公认的哲学术语来指称它。但当我们说一个事实依据另一个事实而确立，或者说一个事实使得另一个事实被确立时，我们有很多熟悉的语词可用来点出其正确指向。举一些例子可能会有帮助：

析取性（disjunctive）事实基于其真实的析取支。我现在要么在新泽西，要么在剑桥，这便是一个析取性事实。这个事实依据“我在新泽西州”这一事实（析取支）而确立（当它发生时）。如果我已经在剑桥，那么尽管原因不同，但我们会得到相同的事实。

存在性（existential）事实凭借其具体事例而确立。有人打翻了牛奶，这是一个事实。譬如，当弗雷德打翻牛奶时，我们便得到了“有人打翻了牛奶”这个事实。如果是别人打翻了牛奶，那么尽管原因不同，但我们会得到相同的事实。

有关具有可确定（determinable）特征的事物的事实以更多的确定性事实为基础。“某个球是红色的”这个事实依据（譬如说）深红色的球而得到确立。“粒子的质量在10～20 MeV之间”这一事实依据粒子的质量为17.656 MeV这一事实而得到确立。

涉及可定义的（definable）属性和关系的事实依据其“定义的扩展”而确立。根据定义，所谓正方形是指一个等边的矩形。鉴于此，如果ABCD是一个正方形，那么“它是正方形”这一事实就依据它既等边又是矩形这个事实。后者便是使一个正方形得以确立的东西。

伴随性（supervenient）事实典型地依据它们所伴随的事实，尽管我们不能有系统地说明依赖的模式。“2008年美国对中国的贸易赤字大约为1170亿美元”这个事实伴随着大量单个经济交易活动的事实，也许最终还可以说伴随着有关（构成从事买卖的人的）夸克和电子等一系列更广泛的事实。宏观经济的事实正是依据这些较低一级的事实而确立，尽管我们不可能具体找出，哪怕是原则上找出，是哪些微观事实使得这一宏观事实得以确立。

这些实例给出了支撑关系的两个重要特征。首先，这种关系是必然性的一种形式。支撑一个给定事实的事实需要其所支撑的事实具有绝对必然性。这一点将支撑关系与某种形式的因果关系或推理确定的关系区分开来。毫无疑问，从某种意义上说，效果取决于其原因：使效果得以显现等的原因。但正如休谟指出的，总是存在这种可能：原因发生而不产生效果。正如上述例子所标明的那样，我们感兴趣的支撑关系涉及一种更为内在的依赖形式。

其次，支撑关系是一种解释性关系。引用一个支撑给定事实的事实是为了给出有关为什么能得到这个（给定）事实的信息。因此，提出支撑关系是一种客观关系的假设，正如我做的这样，就是假设存在有关解释性命令的客观事实（这不是说给出解释的实践始终只是一件报告这些客观事实的事情）。

这些论述并不等于给出支撑关系的定义。我自己的观点是，这种关系太过基本以至于我们给不出定义。对它的理解更多的是通过非正式的解释，但我希望这里的论述已足以满足我们目前的需要。[75]

建议

在数学哲学上还原论者的建议是这样一个断言：某一数学分支的每一项数学事实——例如算术的每一条真理，或集合论的每一条真理——最终都基于不同种类的事实，例如关于形式的可证明性的事实，或关于

如果存在无穷序列对象时所发生情况下的事实。[76] 我们已经考虑了两个例子，其中还原论者的这类建议似乎暗示着算术对象并不那么真实或像某些其他事情那么“像回事儿”。在这些建议中，还原关系已经有鲜明的特征：较高层次理论的对象并不包括理论被还原到的更为基本的事实。这暗示了一种用于解释有条件实在论的形而上学论点的自然战略。

让我们这么说：一个事实是基本的，如果它没有进一步的事实作为支撑的话；一个东西是基本的，如果它是基本事实的一个组成部分的话。由此，我们可能用“某些数学对象是基本的东西”这一观点来确认关于数学的十分牢固的实在论。它肯定是铁杆柏拉图主义者的观点。在他们看来，数是独特的、抽象的物质——排列在柏拉图天堂上的不可见的光球。同时它也是那些较温和的柏拉图主义者所持的观点，这些人将算术，也许还包括像策梅罗—弗兰克尔集合论这样的更全面的理论，看作一整套自足的、并非基于更基本事实的真理。这两种数学实在论都反对我们前面讨论的有条件实在论的形式。

更一般地说，用来解释与有条件的实在论相联系的令人费解的空话的建议可陈述如下：

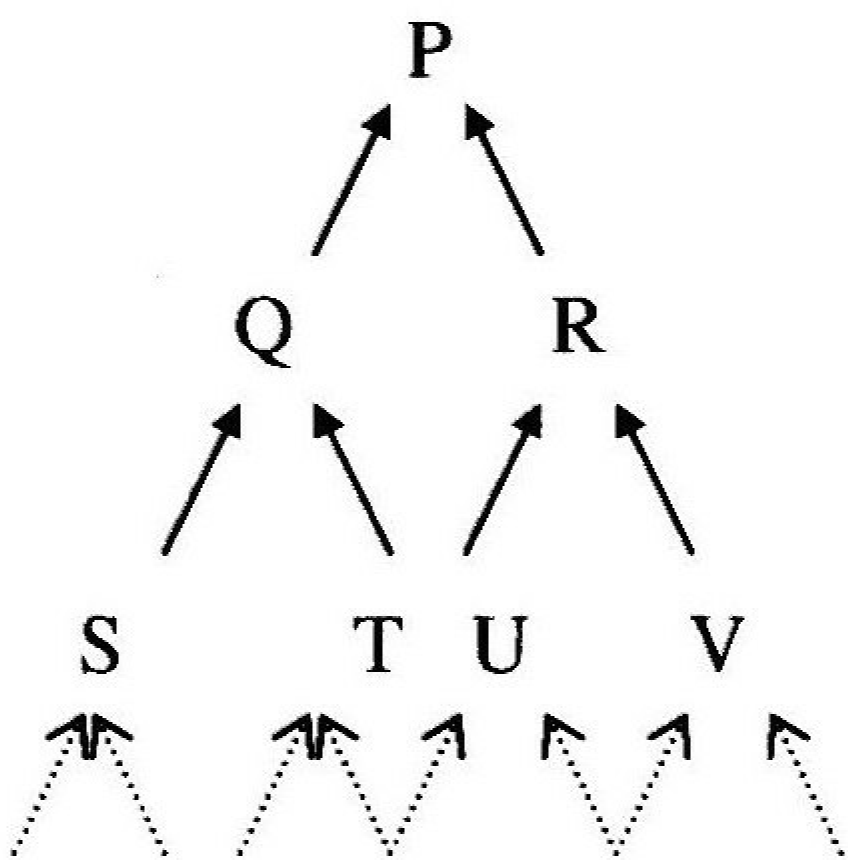
有条件的实在论关于F的观点是F存在，且没有基本事实将F作为其组成部分包含在内。

当哲学家告诉你，数在形而上的意义上不是真实的（即使它们在数学的意义上是存在的），那么他或她的意思可能是指这样一件事：数所代表的每一个事实最终都将凭借一些数不作为其组分的事实的集合而得到确立。

这种观念得到了来自对其他领域“突现的 [77] ”或“更高层级的”实体的本体论地位反思的支持。譬如以“美元”“欧盟”或“我的硬盘驱动器里的编码信息”为例。一望便知，这些东西是存在的。美元作为真正的国际流通形式是一种存在，在这个意义上，意大利里拉也曾经存在，但现在已不再存在。欧盟是一种存在，在这个意义上，星际联邦不存在，但也许有一天会存在。我的硬盘驱动器里确实编存了一大堆信息，如果我按错键，这些信息就可能会被破坏。然而，有很强的倾向认为，尽管对这些存在的断言是完全正确的，但将美元看成是像桌子或上帝一样的东西，或是像柏拉图主义者眼中的数一样的东西，则是错误的。上述建议

使我们能够理解这一趋势。作这样的假设是很自然的：有关美元或欧盟的每一个事实最终都是基于如下一些不同的事实——经济行为者和诺克斯堡的黄金储备局的态度和行动，和欧盟国家间做出的法律安排等事实。当然有一点希望在此说明一下，美元对欧元的比价目前是0.6825欧元。但这一点不是建议所要求的。与还原论的一些早期概念不同，我们这里所理解的还原论不是一种关于用更高级词汇来阐述句子含义，或有关将这样的句子翻译成另一种语言的可能性的观点。它只是这样一种断言：有关货币和其他一些事情的每一个事实最终都是基于关于其他事情的模式极为复杂的事实。当然这种货币还原论未必是真实的。这在经济哲学里是一个实质性问题。上述建议只是说，我们的直观感觉是，从本体论角度看，美元是“派生的”或“第二位的”。我们的这种感觉源自我们的强烈质疑：从上述意义上看，美元不太可能是基本的东西。

如所制定的那样，这个建议要求仅当存在基本事实时才存在真正的东西。人们可能会反对说，我们不能先验地假设这一点。但我们都知道，事物的层级或树状结构可能是无限的，其中P基于Q和R，而Q和R又是基于S, T，U和V，如此循环往复。



作为一个粗略模型，我们可以想象这样一个世界，其中关于原子的事实是建立在夸克和电子的事实基础之上，而夸克和电子的存在又是建立在“超夸克”和“超电子”的事实基础上，如此以至无穷。当然，仅仅存在一些非支撑性的树状分支并不意味着不存在基本的东西，因为基本事实也可能存在于其他地方。只有当每一个事实都基于进一步的事实时，麻烦才会出现。在这种情况下，我们的定义要求不存在基本的东西。

这是否就是我们的建议受人诟病的后果？解读这一点的最佳理由可表述如下。其中每一个树状分支都有其支撑基础的世界可分为两种。在这些世界里，每一个作为某一层面上事实的组成部分的对象，最终都绝对会在我们沿着树状结构不断向下探索（追溯每一事实的根源性事实）的过程中“消失”。在这种情况下，我们的建议要求，很多事情虽然在通常意义上是存在的，但最终没有任何东西是存在的。而一切事物都将获得如同美元一样的形而上学的地位，这似乎是正确的。如果我们逐渐提高我们的形而上学显微镜的放大倍数，可以看到旧的东西在不断消解，新的事物在不断出现，因此我们似乎可以正确地说，从形而上学上看，一切事物都可以视为一个标杆，没有任何东西具有基本不变事物的最终地位，如果我们认为存在这种事物的话。

但也存在这样一种可能：即使每一个事实都是基于进一步事实，但有些事物却是恒久不变的，在这个意义上说，当x是某个事实[…x…]的组成部分，那么它也是每个以[…x…]为其基础的事实的组成部分。直觉上（反对意见），一个恒久存在的事物应当在每一点上都像基本事物那样真实，并且一个事物可以恒久存在而不具有基本存在的属性，因此我们不应当说一个事物是真实的，当且仅当它是某个基本事实的组成部分，而应当说一个事物是真实的，当且仅当它要么是基本的，要么是恒久的。

对此，我认为这个问题将不会出现。假设x在预期的意义上是恒久的事物，且考虑x是自我认同的事实：[x=x]。我们完全不清楚这样一个简单事实如何可能建立在更基本的事实基础上。由于这个事实就是基本事实，因此x就是一个基本事物。如果这是正确的，那么我们可以保留我们的定义，因为任何恒久的东西也将一定是基本的东西。

综述和存在的突出问题

思考数学的哲学家往往会发现自己被拉向两个方向。一方面，数学就其自身和在科学中的作用而言，取得了巨大成功，加上它的一些基本原理非常明晰，这使得哲学家倾向于假设标准数学的断言必定是真实的，因此各种不同的数学对象必定存在；另一方面，许多哲学家认为数学对象在形而上学方面等同于范式（我们已经给了一些具体情形作为例子，但我们也可以考虑上帝和天使，如果你认为他们存在的话）的想法十分荒谬。我们一直在努力确立一种立场，以便对这些倾向做充分公正的评判，为此我们有一个建议。关于数学对象的有条件的实在论持这样一种观点：虽然数及其类似的概念在肯定其存在性的句子字面上没错的意义上是存在的，但它们在下述意义上不是基本的：出现数的每一项事实最终都是建立在不含数的成分的事实基础上。

顺便提一下，我们也许注意到，这种说明解释了为什么数学上的某些唯心主义理论和建构主义理论常常被视为与十分牢固的实在论不相容。数学唯心主义者（这里用一个标签来涵盖这种哲学的各种不同的观点）同意，至少标准数学的某些核心断言是对的，但他们坚持认为，数学事实在某种程度上是建立在我们的数学思想或活动的基础上的。这种认识的最生猛的版本将数学对象等同于大脑中的观念（但是谁的大脑？）。而较微妙的版本则坚持认为数学真理以某种方式基于我们的实践，并坚持认为之所以如此只是因为我们已经接受了数学框架，在这种框架下，算术计算的一般规则被认为是有效的，（譬如）我们可以肯定地说，235+657=892。这种观点的支持者对于我们是如何通过实践使得数学对象变得存在的问题往往是守口如瓶。很明显，一直以来我们并不是像建筑师造房子那样来构建数的。但如果这种比喻是条死胡同，那么在什么意义上我们可以说数在某种程度上是我们创造？我不想继续追究这个问题，只想说，我们的纲领提供了一种陈述这一立场的方法。数学唯心主义的总体方案如下：

对于每一个数学事实[A]，

[A]凭借……（有意识生命的思想、活动或实践）得以确立。

事实上，这是唯心主义和建构主义在哲学各领域通行的一个好的一般性纲领。贝克莱的唯心主义认为，外部对象可以按这样一种断言来理解：包括像桌子和椅子这样的每一个事实最终都是基于有关上帝的想法的事实。穆勒的世俗唯心主义则用有关人类感觉的事实取代了有关上帝

的事实。康德的道德建构主义理论可以表述为这样的观点：关于对或错的每一个事实最终都基于判断的事实，任何理性的代理人在考虑该怎么办时，都会诉诸这种判断。一度在人文领域很是时尚的社会建构主义学说可能持这样一种观点：有关社会实在的事实（在荒谬的极限情形下，直接就是关于实在的事实）最终都基于我们在接受或拒绝关于这种实在的断言方面的认知实践上的事实。这些观点勾勒得非常粗略，但它们都具有正确的一般形式：一种事实——那种似乎不直接关联到人的思想的事实——被认为是建立在关于人类的思想或实践的事实基础上的。当我们考察基本事实时，更高级别言谈的独特对象——普通对象、社会阶层、道义责任等——都消失了，取而代之的是不同类型的项目：陈述、观念、人及其实践等。这些唯心主义的激进形式可能会被认为是错的，但它们对于哲学家和其他领域中具有哲学头脑的学者来说则具有长期的（和神秘的）吸引力，因此值得我们在此对它们寻求一个明确的说法，并解释为什么我们会认为它们与无条件的实在论是不相容的。目前的框架提供了这样一种解释的开端。[78]

现在，让我们回到数学的情形。我们的纲领使我们出于一种接受如下形式的还原论观点的立场：

对于某个领域的每一项数学事实[A]，

[A]凭借……（不将数学对象作为其组成部分的某种事实）而得以确立。

但接受了这样的观点，我们该如何来评判它们呢？随着我们的例子开始展现，我们看到，有很多种方法将算术事实与更基本的事实——证明理论的事实、集合论的事实、纯模态逻辑的事实等——联系在一起。在这些更基本的事实中，算术对象不作为其组成部分出现。这些建议都有足够的材料支撑：它们真理与真理配对，并且经过适当的构建，它们保留了逻辑关系。事实上，在许多情形下，它们遵循更严格的（虽然有点难以捉摸）约束，我们可以称之为相关性。如所指出，对算术陈述A的任何证明通过琐碎的步骤很容易转换为对其相应的形式主义或模态结构主义的证明，反之亦然。严格来说，“在15和20之间存在两个素数”的事实与“每一种可能的ω序列在第16个元素与第21个元素之间存在两个素数”的事实，或与“句子‘存在两个素数……’在PA_{ω}中是可证明的”的事实，是有差别的。这些事实之间之所以存在差别是因为它们有不同的成

分。但这些断言的数学意义是如此接近，使得我们很想说，从数学的角度来看，它们代表了同一个事实。我不这么说，因为我不知道该如何解释这些断言所涉及的事实中同一的概念。（我自己的解释过于细致无法支持这种主张）。不过，人们可能会说，任何像样的算术还原论必须具有这样的特点：由算术事实还原得到的事实至少必须具有与算术上原初的意义大致相同的数学意义。令人担忧的是，即使我们施加了这个约束，在算术上还是有许多同样引人注目的还原论建议。一般来说，对于数学某个领域的问题，如果有一个合理的还原论解释，就会呼啦一下出现许多个这类建议。[79]

这一点之所以令人担心，是因为人们会很自然地认为这些建议不可能都是正确的。如果一个有关的数的事实凭借句子在PA_{ω}中的可证明性的某个事实而得到确立，那么就很难相信它还能通过有关所有ω序列的完全不同的事实而得到确立。这里应当强调的是，这类过分的决定性原则上是可能的。如果我们有一个析取性事实[p或q]，其中p和q都是真实的，那么[p或q]要么凭借[p]而得到，要么凭借[q]而得到。但这样一种想法——事实（算术事实）的统一域可能会被源自两个或多个不同域的事实大规模地和有系统地过度决定——似乎很难令人置信。然而，如果我们拒绝这种可能性，那么还原论的观点就是必然的选择。对这种选择他可能凭借什么样的基础呢？这很难说。

对这种困境的一种反应是采取一种怀疑论的形式。如果竞争性的还原论方法是不兼容的，如果我们没有在其中进行选择的基础，那么可以肯定，唯一适当的反应就是暂停判断。根据这种观点，关于数学事实是否建立在以及如何建立在更基本的事实基础之上的问题是完全有意义的，但也是无法给出最终答案的。这不是一个荒谬的想法。为什么我们一定要利用各种资源来回答我们能接受的每一个深奥的形而上学问题？不过，知道我们是否能抵御这种令人失望的结局这也不错。

这里有一条可能的出路。到目前为止，我们都是自上而下地考虑问题：我们从数学事实开始，并寻求这些事实赖以确立的更基本的事实。这种探索的前提是，当我们考虑一个普通的数学断言——例如，断言在15和20之间存在两个素数时，我们有一个明确的事实，一个我们可能会考虑其基础的事实。但这也许是一个错误。假设事实证明，对于算术的每一个推定的基础——形式主义基础、模态结构主义的基础等——都存在不同的数域，那么关于数的事实便是由该问题的基本事实构成的。根

据这种观点，根本就不存在自然数系这样的东西。存在的只有形式主义的数，关于这种数的事实则是基于有关PA_ω内的可证明性的事实；还存在模态结构主义的数，而关于这种数的事实则是基于有关所有可能的欧米茄系统的事实，等等。由于这些数系之间的差异不产生任何数学上的差异，因此数学的语言和实践没有办法来区分它们。如果这是正确的，那么普通的数学语言将充满了语义不确定性。当我指着远处的一条河说："我们把这条河叫作'奥夫'吧。"我引入了一个有意义的词，但因为我懒得费心确定指的是附近那么多像河一般的对象中的哪一个对象，所以我的这个称呼并不指称具体哪一个东西。相反，它"把它的指称分发给了"一定范围内的所有候选对象——有些对象比另一些对象宽一点，有些则是长一点，等等。当我在一个句子里使用这个词的时候，我的脑海里根本就没有确定要指称这些候选对象中的哪一个。在这种情形下，谈论奥夫是世外桃源这个事实就是一种误导。由于附近有许多河流状对象，有很多这样的事实，它们之间的差异对我的目的来说并不重要。同样的，如果存在很多类似数的对象的系统，它们之间的差别只有通过这些数系所基于的事实，而不是通过任何数学上的重要特征，才能被区分开，那么谈论235+657=892这个事实就没有意义，因为附近有许多同等条件下的事实，每个事实涉及某种确定类型的数，每个事实都以某种确定的方式基于一些基本事实。[80]

如果说这是一种语义形而上学的情形，那么就难怪我们不知道如何回答有关算术事实最终如何得到支撑的问题。我们不知道该如何回答这些问题，因为它们有一个错误的前提，即对于数学对象和有关它们的事实（普通的算术语言设法指称）的确定范围，存在一个独特的系统。

对这种形而上学多元论和语义不确定性之间的结合最终是否协调还远远没有定论。为了解决这个问题，我们需要一种有关一种事实在何种情况下会导致或产生另一种事实的条件的一般性理论。产生这样一种一般性理论的任务是十分艰巨的。但让我们暂且假设这种观点不仅逻辑上前后一致，而且是正确的。于是我们会注意到，即使有关数的问题因提法欠佳而被拒绝，我们仍可能认可关于算术的有条件实在论的形式。因为我们有可能证明，算术语言对每个候选对象的解释采用的是一种用来描述一类因还原而消失的对象的语言。如果是这样的话，那么说"算术对象不会出现在基本事物中，因此算术对象最终不是实际存在的"这句话就仍然是对的。

这给我们出了一道在此框架内最难啃的难题。假设每一个推定的还原作用都如上述那样与其自身的“数”类相联系。这样虽然消解了不同竞争性还原作用之间的矛盾，但在认为“每个数系都可按这种方式被还原”的有条件的实在论与认为“在基础水平上至少存在一个类似数系的项目”的十分牢固的实在论之间，仍然存在一种理解上的分歧。在我们一直讨论的框架内，这些都是有意义的假设。但还没有什么办法能够决定哪些假设是正确的。

在这一点上，某些哲学家会倾向于采用简约原则。他们会说，在对基本实在的解释方面，我们应该将结论建立在尽可能少的几件事情（或几类东西）上。如果数学哲学并不需要存在基础水平的数——如果可还原的数永远“起作用”——那么基础水平的数就是可去除的，我们的理论家应该拒绝它们。以这种方式进行便是先验地假设了基本实在是一种非常稀罕的晴空美景和沙漠景观。就我个人而言，我认为这个假设是完全没有根据的。但是，如果简约化无法帮助我们，我们可能会发现自己处于一种进退两难的境地，唯一合理的反应是对判断的悬置。换句话说，我们可能会发现自己被推向这样一种观点：认为有条件的实在论关于算术所提出的问题是完全有意义的，但我们想象不出有什么方法可以给予解答。

在哲学的某些方面，这种僵局表明我们的问题提得很不恰当。现在这种情况是不是就是这样呢？我不认为如此。我相信，出现在这种争论中的支撑关系是（或能够变得）完全可以理解的。因此，数学实在论的难题，即某些数学对象是否在基本层面上存在的问题，是清楚的。我得承认我不知道如何去回答。但这并不是说，这个问题最终无法回答。哲学家们并不总是能够将普通的本体论问题（即关于什么东西是普遍存在的问题）与深刻的本体论问题（即关于什么东西在基础层面上存在的问题）区分开。因此，我们还没有明确范式来调查后一类问题，也不清楚拿什么来建立关于什么是“最终实在”的断言。既然我们已经做了区分，那么我们就能够查看记录，以便探询我们是否在哲学上或物理学上甚至神学上有具体指向有关基本实在的断言的例证。如果我们找到了合理的例子，我们便可以从它们那里得出明确的方法论启迪。研究数学形而上学的方法可能包括这些与数学真理的基础有关的方法论原则。这项工程能见效吗？在这一点上，我觉得我们还没有明确的答案基础。

评吉迪恩·罗森的“数学对象的实在性”

蒂莫西·高尔斯

在吉迪恩·罗森这篇文章的开头，他描述了关于数学的某种哲学立场。自从我自己一度曾思考有关数学的各种哲学观点以来，这些观点一直是那种样子。我不是一个专业的哲学家，我从来不试图制定出一个全面详细的方案来维护我的观点，但我一直相信，它们是基本正确的。

如果说有人能够动摇这样一种自信，那么这人便是罗森，一个具有非凡的表达能力和仔细分析他人哲学立场的人。他还有一套提出观点的方法，直觉上我对那种关于数学对象的实在论观点有反感，但他提出问题的方式使我的反感明显减弱。举例来说，如果你对罗森说，你不相信那种认为数是飘浮的并具有各种复杂的相互关系的形而上学论调，他会告诉你他也不相信。然后他可能会问你，为什么你说存在无穷多个素数，如果你实际上并不相信是这样的话。经过这种短暂的交谈，你会变得很难准确把握在非实在论观点与罗森所持的精致实在论观点之间到底有什么不同（至少对我来说是这样）。

现在，在我读了罗森的这篇文章之后，对我来说，这些问题都得到了很大程度的澄清。他把他所讨论的立场称为有条件的实在论：大致说来，所谓有条件的实在论，就是认为数学对象确实存在，但正如他所说的，它们在形而上学上处于“第二位”。他本文的目的是要说清楚这可能意味着什么。为了这一目标，他引入了事实之间的“支撑关系”。这种关系粗略地讲就是，如果事实A是建立在事实B的基础之上，那么事实B就比事实A更基本，并且足以解释A。（罗森给出了几个不同的例子来说明和澄清这个概念。）随后他建议，无条件的实在论者与有条件的实在论者之间的区别，就在于前者认为某等些数学事实是基本的（也就是说，没有任何进一步的事实为其基础），而有条件的实在论者认为，虽然数学事实是客观真实，但它们最终都基于非数学的事实。正如他所说的那样：

> 当哲学家告诉你，数在形而上的意义上不是真实的（即使它们在数学的意义上是存在的），那么他或她的意思可能是指这样一件事：数所代表的每一个事实最终都将凭借一些数不作为其组分的事实的集合而得到确立。

罗森说他本人并不认同这个观点，然而此观点与我之间以一种他所希望的方式产生了共鸣：我认为，如果我尝试为我自己在关于数的问题上的有条件的实在论观点进行全面的辩护，那么我的确会这样去做：尝试去证明关于数的事实都是建立在其他事实的基础之上。要实现这一点有几种方法，它们恰好对应于有条件实在论旗帜下的几种不同的哲学立场。罗森详细讨论了其中的两种（这两种都不是我自己可能遵循的方法），同时还一般性地讨论了这种立场的一致性。

罗森的建议的优点之一是，它用“是否所有的数学事实都基于其他事实”这样一种紧密关联而且非常明晰的问题取代了有关数和数学陈述的实在性那样一些模糊不清的问题（说它模糊不清是因为这类问题往往不能明确说明“实在”指的是什么）。回答这个问题是一项很艰巨的工程，但它也明确规定了哲学家，甚至数学家，可以为此做些什么。

第10章　我们从数学中得到的要比赋予它的多

马克·施泰纳

在《皇帝新脑》（Penrose，1989）和最近的《通向实在之路》（Penrose，2005）两本著作中，彭罗斯教授倡导一种他所谓的柏拉图主义，即柏拉图的数学世界。由于彭罗斯教授将出席本次研讨会，因此我想在这里讨论他的这一观点——或更确切地说，是我基于他的这两本著作的强大影响力而认为的他的这一观点——将是适当的。

正如彭罗斯在《通向实在之路》一书中所解释的，他所追求的是数学的客观性[81]，而不是“数学对象”的存在性。关于后面这一点，正像本次研讨会的另一位发言者罗森教授所指出的那样，是指数学是否是一种“没有对象”的学科。正如伯吉斯和罗森在他们的优秀著作（Burgess and Rosen，1997）中指出的，柏拉图主义被奎因挟持了，他把它定义为“对数学对象的量化”，它不再具有它过去所具有的意义。由学者撰写的讨论这方面的书一本接一本（其中大部分由牛津大学出版社出版），例如，有本书就提出，“数学对象”是否应该突然消失，那样它将带来可观的变化。[82]

因此我认为，完全不涉及柏拉图，讨论概念而不是对象，并将笛卡尔作为彭罗斯的认识（譬如认为像曼德布罗特集这样的概念是客观的）的真正来源，这样可能会更好。毕竟，彭罗斯并不（像柏拉图那样）认为每一个概念都是客观的，这一点与笛卡尔非常一致。笛卡尔在《第一哲学沉思集》中曾这样描述他对这个问题的看法：

> ……我们必须注意那些不包含真实的和不变的本性，但仅仅包含由理智虚构和组合成的观念。这些观念总是能被同一个理智分开，不是单单通过抽象，而是通过一种清楚和分明的理智活动，因此凡是不能以这样的方式被理智分开的观念，很明显就不是通过理智组合而成的。例如，当我想象一匹带翅膀的马，或者一头现实存在的狮子，或者内接于正方形的一个三角形时，我很容易领会到，

> 我也可以想象一匹没有翅膀的马，一头不存在的狮子，一个脱离正方形的三角形，因此这些东西没有真实的和不变的本性。但是，如果我想象一个三角形或一个正方形……那么任何我理解为包含在三角形的观念里的东西——例如其三个角度之和等于两个直角——我都可以真实地断定它是属于三角形的…… [83]

笛卡尔在这里说的是，你从数学里得到的要比你赋予它的多。数学思想中内在地具有一种没包含在语词定义中的“潜在信息”。[84] 这种“潜在信息”是这些思想精髓的一部分，而不是由数学家放进去的。在我看来，这一点正是数学与游戏之间的真正差别。不妨看看下面这个国际象棋残局（见图10.1）。[85]

假设双方都有完美的发挥，那么白棋要准确地将死对方需要走262步。而在许多局面下，走的那些着“没有任何意义”，这里的意义是指在不实际摆出所有可能的着法谱图的情形下我们不可能解释清楚这个结论。虽然这种事情显然不是“预期的”（国际象棋里有50步判和规则，是指在终局前双方都没有走动任何一兵，也没有吃过任何一子，即判和），这里出人意料的是，国际象棋规则所包含的条款非常之少，你怎么一点收获的投资。因此，国际象棋的本质就是其着法的任意性，这肯定不是笛卡尔所说的那种“真实的和不变的本性”。

我们不妨将上述情形与彭罗斯教授喜欢列举的“神奇的”复数的例子相对照。当意大利人将复数作为方程的实数解之外的虚数解引入时，没人能预言它们在与实数的关系上会扮演什么角色。我们来看看下面这个优美的方程

$$e^{i\pi}+1=0$$

它是欧拉发现的著名公式

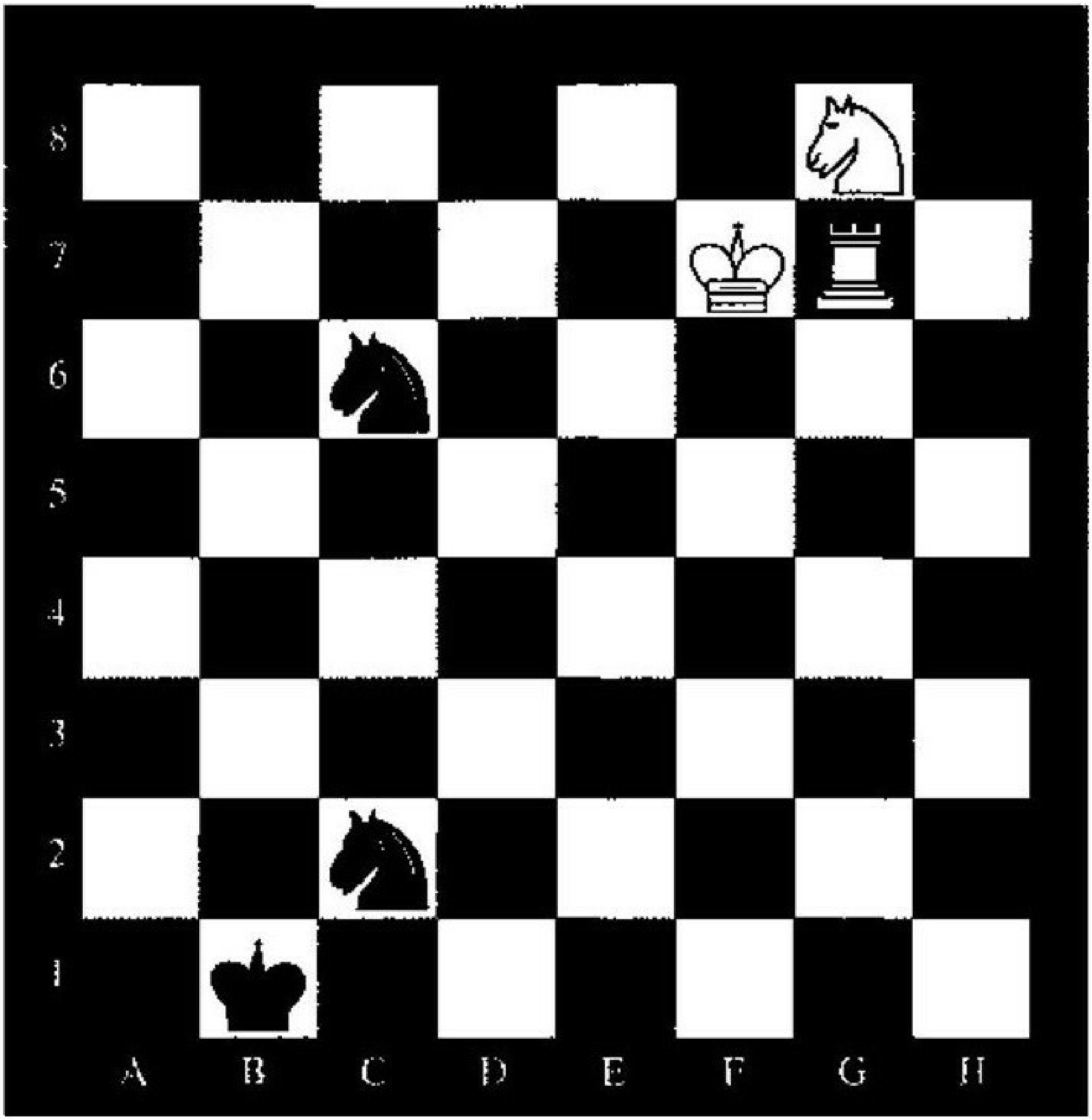


图10.1轮白棋走，262步将死对方

$$e^{i\theta}=\cos(\theta)+i\sin(\theta)$$

的一种特殊情形。

我们注意到，当引入虚数后，将实数提升为虚指数的思想即使对于像卡尔达诺和邦贝利（Bombelli）这样的数学大家来说也是不可想象的。然而，一旦人们认真领会了这种思想，在如何处理上就很少有或者

说根本没有选择。

虚数的最初提出者没有认识到的虚数的另一个属性是复数的绝对值。当我们将复数用欧几里得平面来表示时，这个问题便显现出来了。利用这个属性，数学家可以解释有关实数的事实，例如，为什么在实轴上处处有定义的实函数1/（$1+x^2$ ）不等于其幂级数展开式$1-x^2 +x^4 -\ldots$，其中$|x|\geq 1$（绝对值等于1的复数组成围绕原点的一个圆，实数1和虚数i都在这个圆上。对于i，这个函数没有定义，尽管它在复平面上连续，因为这里的分母是零。复分析里的标准定理能够解答余下的问题）。

有人可能会说，虚数的引入部分是出于计算上的方便。卡尔达诺甚至用虚数来计算一元三次方程的实根（见其著名公式）。但即便如此，其思想中所固有的“潜在信息”也远远超出了方便计算的范畴。总之，引入虚数的意义从《皇帝新脑》的下面这段话可见一斑（这段话就像是笛卡尔写的，都说他能预言数学的未来）：

> 初看起来，引入负数的平方根似乎只是作为一种工具——旨在实现特定目的的一种数学发明，但后来人们看得越来越清楚，这些对象所能实现的作用远远超过了它们最初的设计。正如我在前面提到的，虽然引入复数的初衷是为了使平方根可以畅行无阻，但后来人们发现，通过引入这种数，像是奖励，我们可以求得任何其他形式的根，或求解任何代数方程。后来我们还发现了这些复数所具有的其他多种神奇的属性。对于这些属性，我们起初没有感觉到一点征兆。这些属性原本就在那里。它们不是由卡尔达诺放进去的，也不是由邦贝利、沃利斯、科茨、欧拉、韦塞尔、高斯和其他伟大的数学家放进去的，尽管毋庸置疑这些人都富于远见卓识。人们逐渐发现，这种神奇的性质是结构本身所固有的。卡尔达诺在引入复数时，对这种数所带来的许多神奇特性没有一丁点思想准备——这才有后来的这些属性以不同的人来命名，例如柯西积分公式、黎曼映射定理以及莱维延拓性质等。这些以及许多其他显著的事实，正是卡尔达诺在1539年前后第一次遇到的没有附加任何修正的那种数的性质。

彭罗斯（1989年版，第96～97页）[86]

笛卡尔/彭罗斯关于数学概念里“真实的和不变的本性”的思想，我在

几年前的一篇文章中对此作过讨论（Steiner，2000），与数学应用于自然没有特别的关系，但它是本论文的立论所在。人们常常看到，正是这种在数学概念发现的潜在的数学信息——甚至它们的引入是出于“方便”的考虑——提供了数学在自然科学领域的最壮观的应用。这个“剩余价值”在“虚”数应用到“真实”自然方面尤为明显。正如在纯数学中的情形，这些应用最开始是出于计算上的方便考虑，但终了它们具有了一种描述上的必然性。在下文中，我将讨论一些非常有名的事实（读过彭罗斯著作的人都了解这些）。人们对这些事实是如此熟悉，以至于可能忘记了它们有多么了不起。从现在起，我们将讨论数学在自然科学中的那种所谓“不可思议的有效性”。

欧拉的发现使我们可以很方便地表示平面上的转动——即通过复平面上单位向量——而不必借助于凌乱的三角公式。两个转动的组合可以用两个单位向量的乘积给出，结果仍是一个单位向量，其辐角等于相乘的两个单位向量的辐角之和。

19世纪里有许多这类将这种方便推广到空间转动的尝试。欧拉曾证明如何用三个角（现今称为“欧拉角”）来表示空间转动，将复数概念推广到“复空间”上的三维向量似乎是合理的，但这些尝试都失败了。于是哈密顿通过与复数类比，不得不将空间维数提升到四维，以便得到一个满足乘法运算的向量空间。[87] 他将这种代数的元素称为“四元数”，写成a+bi+cj+dk的形式，其中单位元素的乘法满足$i^2=j^2=k^2=ijk=-1$，他曾将这个方程刻在一座桥上，以铭记他发现这些关系所带来的激动之情。与复数类比可见，单位四元数表示空间转动。乘以两个单位四元数，表示两种基本转动的组合，由此我们有（用后来的术语）一种同态。四元数乘法不满足可交换性，这一事实对于这里的目的是微妙的，因为转动本身就是不可交换的。然而，这里还有一些令人困惑的地方：在i、j、k分别表示关于x、y、z轴转动的特殊情形下，这种转动是转过180°而不是90°，否则由i、j、k所表示的连续转动不会使轴回到原初位置。（更仔细的分析见彭罗斯《通向实在之路》2005年第1版，第11章）。这意味着三个单位四元数-i、-j、-k分别表示转动540°（=180°）。一般来说，这是对的——负的单位四元数表示与四元数本身相同的转动。同态不是同构，但两个同态等于一个同构。关于固定轴的空间转动的任何连续（子）群r（θ）都可以通过单位四元数的连续路径q（θ）同态地表示出来，只不过q（θ+2π）=-q（θ），而r（θ+2π）=r（θ）。每次转动确立两个标签，或“宇称”。转动的宇称是无用的信息，或者说看起来如此。

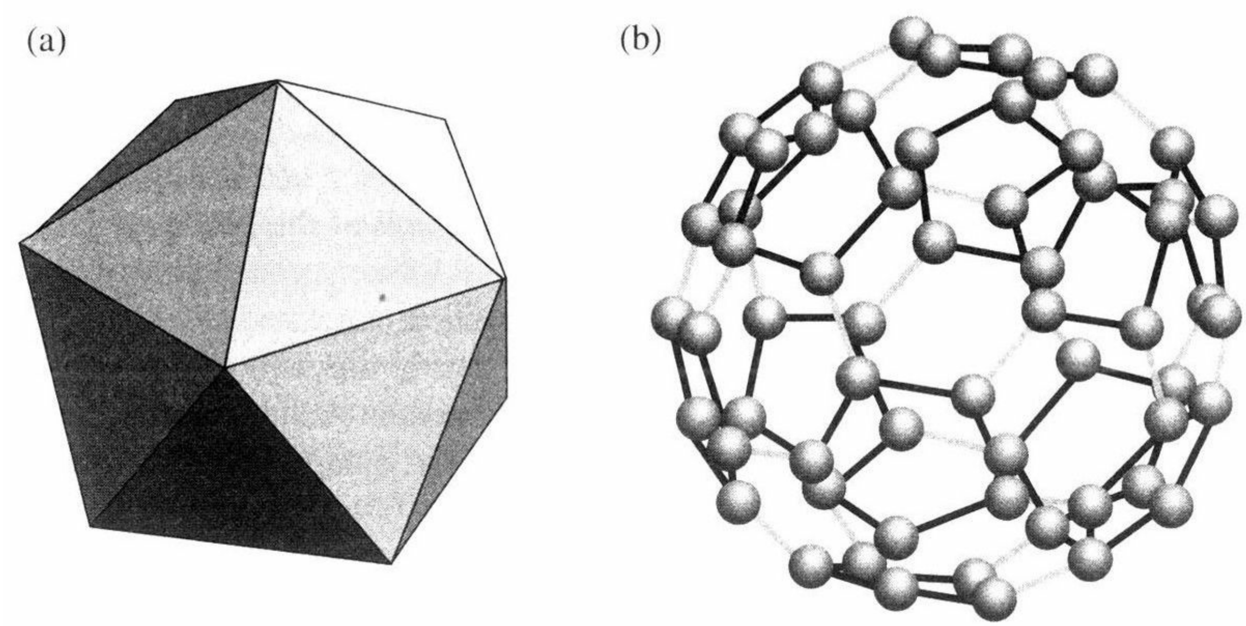


图10.2（a）正二十面体；（b）碳-60分子（巴基球）

将模1的复数的平面转动表示扩展到空间转动的另一种方法出现在与上述方法大致相同的时期。这是一种由（普通）复数的2×2酉矩阵表示的转动。酉阵M满足MM* =I（单位矩阵），其中M* 是通过转置M中的行和列，再将相应的元素用M的共轭复数来取代（即，x+iy→x-iy）所得到的矩阵。酉矩阵的行列式必须是一个绝对值等于1的复数。通过限定矩阵的行列式等于+1而得到的矩阵构成了我们现今所称的SU（2）群，即2×2矩阵的“特殊酉群”。这种方法是凯莱、拉盖尔和其他一些数学家努力的结果。但我能找到的明确给出SU（2）群与转动群之间同态的最早记录是在1884年，当时费利克斯·克莱因（Klein，1888，第34页及其后几页）做过关于正二十面体群的讲座，[88] 并在两年后将同态方法用于他关于陀螺仪转动研究的计算中（见Klein，1922）。

我们可以证明，SU（2）与单位四元数同构，因此也存在二到一地从SU（2）同态到所有的转动。例如，SU（2）矩阵

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

对应于关于z轴的2θ转动。[89] 当θ=0时，矩阵为单位矩阵$I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$，对应

于零转动；当θ趋向π时，矩阵为-I，转动为满转动。两个满转动使矩阵回到I。我们再一次得到一种无端的转动标记来表示转动的正或负，用SU（2）得到在区间（360°，720°）上的转动，计算上的方便是第一时间引入SU（2）的理由。

非常明显的是，这种多余的信息对于理解我们的宇宙的某些基本特性至为关键。电子的自旋对称性恰恰就是SU（2）转动对称。当电子转过360°，所得到的态相当于其量子力学描述（波函数）乘以-1。为了让电子回到其初态，就必须让它转两圈。如果数学物理学家有更好的计算器，他们可能就错过了科学史上的一个重大发现。[90] 这只是人类局限性的诸多实例中的一个例子，这些局限性阻碍的远不只仅是科学的进步。[91]

电子生活在两维的复向量空间并具有SU（2）对称性这一事实甚至在宏观水平上都可以检测到。考虑碳—60分子，它具有巴基球的形状，或曰截去棱角的正二十面体，它像一个足球一样是由多个六边形和五边形组成的。巴基球似乎有正二十面体的I型对称性：绕其中心有60种不同的转动，构成巴基球不变量。[92] 然而，如果我们在磁场中研究碳—60分子的顺磁性行为，这个磁场也具有正二十面体对称性，那么我们就必须“拉回”到使SU（2）的元对应于I的元。于是，巴基球的真实对称群是SU（2）的有120个元素的子群G。这里SU（2）二到一地同态到I（Chung et al.，1994，§9）。

这个故事里还有个故事。E.N.拉盖尔在1867年发表了一封给埃尔米特的“信”。在信中他定义了矩阵乘法，并讨论了它的属性（Laguerre，1867：1898）。在这个讨论中，他引入了矩阵的整数模p，模的概念属于数论的研究范畴，在任何情况下都难有物理应用。如果我们考虑一个2×2的矩阵，设p=5，并将待研究的矩阵限定为行列式为+1的可逆矩阵，那么我们恰好能得到120个矩阵构成一个群，可以证明，这个群同构于G——一个令人惊讶的事实——这意味着拉盖尔的数论思想被证明可以用于描述巴基球的电特性和磁特性。显然，在此我们得到的要比投入的多（Chung et al.，1994，§2）！

现在让我们回到电子本身的自旋问题上来，电子自旋两周后返回到初态。人们可能会认为，由于电子不具有标准的几何形状，因此可能转过3圈或其他数目的圈数后也会返回到初态。然而，如果我们用拓扑来

理解时数字2就显得很“自然”。

人们经常引用这样一个例子：如果你将一条带子的一端夹在一本书里，另一端拿在手上，然后将书转动一整圈，你得到的是一条扭成麻花的带子。但是，如果你把书转动两整圈，然后让麻花状带子绕书转一圈，则纸带恢复到原状。[93] 拓扑上这个过程表示如下：如果我们将转动做成一个闭环，那么这个闭环是不可以“缩”成一个“点”的（这里的“点”是指恒定的转动曲线，即不转动任何东西的曲线）。只有双环可以这么缩成点。这不是欧几里得空间下的一个直接的事实，它是通过研究SU（2）与SO（3）之间的二到一同态性质来揭示的（见Sternberg，1994，§1.6）。

费曼给出了关于这种思想的另一种“扭曲”（见Feynman和Weinberg，1987，第56～59页）。他建议用一条带子拴住两个电子，并让它们交换位置。于是带子被扭曲，拓扑上看就是我们得到了一种单一的转动，每个电子相当于贡献了一半的转动。在这种情况下，我们预期描述两个电子的波函数会变号，这是费米子的基本性质。如果这种说法是对的，它似乎好得不像是真的，那么我们用SU（2）和一点点拓扑知识便得到了著名的“自旋”与“统计”之间的联系——不用相对论，也不用场论。即使这种想法得不出结论，这种将SU（2）对称性与费米统计联系起来的证明，是SU（2）的发明者做梦也想不到的。

最显著的或许要算SU（2）和一般酉矩阵的进一步推广应用。这些应用使得酉矩阵的概念越来越远离其最初的应用——转动的表示。

我举两个例子：第一个是海森伯在1932年将SU（2）对称性应用于核物理。业已证明，中子和质子是同一种粒子（现今称为“核子”）的两种态，从数学上看，这两种粒子可以类比于电子的“自旋向上”和“自旋向下”的两个态。使中子变成质子然后再返回到中子态的“转动”（一种引起粒子的波函数变号的“转动”）不可能看成是物理空间中的转动。不过，核子的SU（2）对称性，即所谓“同位旋”，有经验的结果。这种类比的物理基础，如果存在的话，甚至在今天仍是未知的。

当我们转向高维酉矩阵后，群与转动的简单联系不再成立。例如，SU（3）（行列式为+1的3×3酉矩阵群）就不同态于群SO（n）。它与转动的联系只能通过类比于SU（2）。尽管这样，也许正因为如此，

SU（3）可以描述夸克的三种状态，而核子正是由夸克构成的（因为核子具有整数电荷，而夸克具有分数电荷，这种差别阻碍了对它们的发现）。转动竟然只是更大的冰山的一角。

克莱因在他的演讲（1922，Lecture I）的开头给出了将推广SU（2）二到一同态到转动群的另一种方式。如果我们放弃酉化的要求，只考虑行列式为1的2×2复数矩阵，我们便得到了现今称为SL（2，C）的矩阵群。克莱因证明了（这条定理的陈述见第626页），在SL（2，C）和二维流形上的变换群之间存在一种两到一的同态关系，而这种群被证明同构到正常洛仑兹群L°（具有正行列式并保前向“光锥”的洛伦兹变换群），当我们研究相对论性量子理论时，就会明白这种同态本身有着重大的物理意义。

我认为，“神奇的”复数（彭罗斯语）对转动的适用性及其到SU（2）及其后的推广在数学史上非比寻常，它很好地说明了数学概念往往具有潜在信息的现象，这种信息可以用于数学的发展。更重要的是，大自然似乎利用这些数学上的可能性。

这里有三个要素：数学、自然和人的心灵。这三者中哪一个负责揭示“数学本质”的这种显著的丰富性呢？

致谢

我要感谢什洛莫·斯腾伯格（Shlomo Sternberg）的重要帮助和忠告。还要感谢什穆埃尔·埃利佐尔（Shmuel Elitzur）、希尔万·卡佩尔（Sylvain Cappell）和卡尔·普希（Carl Posy）在多年的教学过程中所给予的有价值的讨论。本项研究得到以色列科学基金会的支持，批准号251/06，我很感谢这种支持。

评马克·施泰纳的“我们从数学中得到的要比赋予它的多”

马库斯·杜·索托伊

得之多于予之的思想是我从事的这个领域最具吸引力的特点之一。我觉得这个比值在数学中可能比在任何其他学科中更大，这也是我为什么选择了数学而放弃了生物学的一个原因，也许还因为我的记忆很糟

糕。一旦你进入这个领域，你只需要一些公理，其他的一切就会奔涌而出。生物和化学似乎总是需要你记住自然呈现的一些东西，这些东西有时似乎是很随意的。

数学家的艺术往往是挑选游戏规则，以便最大限度地放大这个比值。虽然施泰纳在游戏和数学之间做了区分，但我想他在看到以下这些事例后会同样感到满足：围棋“Go”的简单规则带来极其丰富的行棋局面；定义群的三条简单的公理可以带来大魔单群和作为物理学基础的李群。

这让我想起施泰纳的文章所提出的富于挑战性的论点：在抽象的、优美的数学世界与物质的、杂乱的物理学、化学和生物学世界之间存在非凡的协同作用。这也许有点奇怪：如果量子物理学关于世界被量化成离散的碎片的观点是正确的话，那么完美的圆形或直角三角形——数学最基本的对象——可能就没有任何物理实在性可言。我们轻松随意摆弄的无穷大在本质上很可能有限的宇宙中就没有具体的实现对象。但不管怎么说，我们心灵的这些对象到底是如何帮助我们预言混乱的宇宙的未来行为的，这颇令人惊奇。我们创建用来解三次方程的虚数难道与至关重要的描述量子世界的数是同一种数吗？

也许正像施泰纳在回应我的文章时所建议的那样，我们可以有一个拟人化的答案。我们选择我们乐于驾驭的数学。那么数学和带来最初兴奋的源头来自哪里呢？来自对物理世界的描述。埃及人要想知道金字塔的体积，他们需要知道究竟用了多少块砖石。但是要想计算其体积，那就需要发明一种方法，将砖石形状切割成无限多的小块，无限薄的片，这样他们可以重新排列这些石料，使问题变得容易些。于是积分的早期形式便由此产生了。

在实际层面上考虑切割真金字塔这种过程显然是荒谬的。但这样一种从数学世界到我们这个杂乱的真实世界的投射已经建立起来。由于数学世界是以试图描述和预言物理实在来开始其旅程的，因此我们以纯粹抽象的形式所创造的数学，以其固有的内在魅力，会经常发现，这种旅程启程之后自身便被投射到我们这个杂乱的宇宙创生问题中来了就一点也不意外了。

最后一点。有时数学能很好地说明为什么你不能从所投入的对象上

得到更多，从实数扩展到复数再扩展到四元数到八元数，但数学家可以证明，这之后你得不到更多的数的种类。同样，李群E_6 、E_7 和E_8 具有非常优美而强大的结构，但数学可以证明为什么这种结构会到此而止，不存在E_9 。有时候你得到的要比你预料的少。但了解这一点有时像投入少回报多一样令人兴奋。特殊李群之所以特殊，正因为其独特的性质。不过，E_8 能成为构成实在结构的基本粒子模型还是令人惊异的。大自然肯定有好的鉴赏力。

参考文献

Archimedes.The Method .InGreek Mathematical Works .Ⅱ：From Aristarchus to Pappus ，ed.J.Heiberg and trans.I.Thomas，Loeb Classical Library，362.Cambridge，MA：Harvard University Press（1993）pp.221—223.

Arnauld，A.（1964）.The Art of Thinking：Port Royal Logic ，trans.J.Dickoff and P.James.Indianapolis，IL：Bobbs-Merrill.

Benacerraf，P.（1965）.What numbers could not be.Philosophical Review，74 .

Benacerraf，P.（1973）.Mathematical truth.Journal of Philosophy，70 .

Bohm，D.and Hiley，B.J.（1993）.The Undivided Universe .London：Rout-ledge.

Bolzano，B.（1810）.Contributions to a Better-Grounded Presentation of Mathematics .In TheMathematical Works of Bernard Bolzano ，ed.and trans.S.Russ.Oxford：Oxford University Press（2004）.

Borges，J.L.（1941）.The Library of Babel（La Biblioteca de Babel）.InLabyrinths .Harmondsworth：Penguin（1970）.

Burgess，J.P.（1983）.Why I am not a nominalist，Notre Dame Journal of FormalLogic，24： 1.

Burgess，J.and Rosen，G.（1997）.A Subject with no Object： Strategies for Nominalistic Interpretation of Mathematics .Oxford：Oxford University Press.

Changeux, J.-P.and Connes, A. (1995).Conversations on Mind, Matter and Mathematics , ed.and trans.M.B.DeBevoise.Princeton, NJ: Princeton University Press.

Chung, F.and Sternberg, S. (1993).Mathematics and the buckyball.American Scientist , 81: 56—71.

Chung, F., Kostant, B., and Sternberg, S. (1994).Groups and the buckyball.InLietheory and Geometry: In Honor of Bertram Kostant , ed.J.-L.Brylinski, R.Brylinski, V.Guillemin and V.Kac.Boston: Birkh user.

Cicero, M.T.The Orations of Marcus Tullius Cicero .4 vols.London: G.Bell&Sons (1894—1903).

Coffa, A. (1986).From geometry to tolerance: sources of conventionalism in nineteenth-century geometry.InFrom Quarks to Quasars: Philosophical Problems of ModernPhysics .University of Pittsburgh Series, 7.Pittsburgh, PA: Pittsburgh University Press, 3-70.-

Cohen, P.J. (1966).Set Theory and the Continuum Hypothesis .New York: W.A.Benjamin, Inc.

Courant, R.and Robbins, H. (1947).What is Mathematics?An Elementary Approach to Ideas and Methods .Oxford: Oxford University Press (1981).

Curry, H. (1951).Outlines of a Formalist Philosophy of Mathematics .Amsterdam: North-Holland.

Dedekind, R. (1888).Was sind und was sollen die Zahlen .InGesammelte Mathematische Werke, Ⅲ .Braunschweig: Friedrich Vieweg und Sohn.

Detlefsen, M. (2005) Formalism.InThe Oxford Handbook of Philosophy of Mathematics and Logic , ed.S.Shapiro.Oxford:

Oxford University Press, pp.236-317.

Dorr, C.（2008）.There are no abstract objects.InContemporary Debates in Metaphysics ，ed.T.Sider, J.Hawthorne and D.W.Zimmerman.Oxford: Wiley-Blackwell.

du Sautoy, M.（2008）.Finding Moonshine: A Mathematician's Journey Through Symmetry .London: Harper Perennial（2009）.

Dummett, M.（1978）.Truth and Other Enigmas .Cambridge, MA: Harvard University Press.

Feynman, R.P.and Weinberg, S.（1987）.Elementary Particles and the Laws of Physics .Cambridge: Cambridge University Press.

Field, H.（1980）.Science Without Numbers .Princeton, NJ: Princeton University Press.

Field, H.（1984）.Is mathematical knowledge just logical knowledge?Reprinted with a postscript inRealism, Mathematics, and Modality .Oxford: Blackwell（1989）, pp.79-124.

Field, H.（1991）.Metalogic and modality.Philosophical Studies ，62（1）: 1-22.

Frege, G.（1884）.The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the ConceptofNumber ，trans.J.L.Austin, 2nd edn., New York: Harper（1960）; 2nd rev.edn., Evanston, IL: Northwestern University Press（1968）.

Frege, G.（1893）.Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet ，Vol.I Hildesheim: G.Olms Verlag（1962）.

Frege, G.（1903）.Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet ，Vol.Ⅱ.Hildesheim: G.Olms Verlag（1962）.

Gabriel, G.etal ., eds. (1980) .Gottlob Frege: Philosophical and Mathematical Correspondence .Chicago: University of Chicago Press.

Gödel, K. (1947) .What is Cantor's continuum problem?Revised and expanded versionin Kurt Gödel: Collected Works , Vol.Ⅱ.Oxford: Oxford University Press (1990) .

Gödel, K. (1951) .Some basic theorems on the foundations of mathematics and their implications.In Kurt Gödel: Collected Works , Vol.Ⅲ.Oxford: Oxford University Press (1995) .

Godstein, H. (1980) .Classical Mechanics , 2nd edn.Reading, UK: Addison-Wesley.Gregory, R. (1969) .Eye and Brain: The Psychology of Seeing .New York: McGraw Hill.

Hadamard, J. (1954) .The Psychology of Invention in the Mathematical Field .New York: Dover.

Hardy, G.H. (1940) .A Mathematician's Apology .Cambridge: Cambridge University Press (1967) .

Hartle, J.B. (2003) .Gravity: An Introduction to Einstein's General Relativity .San Francisco: Addison-Wesley.

Hellman, G. (1989) .Mathematics Without Numbers .Oxford: Oxford University Press.

Herschel, J. (1841) .Review of Whewell's History of the Inductive Sciences and Philosophy of the Inductive Sciences.Quarterly Review , 68 : 177-238.

Heyting, A. (1931) .The intuitionistic foundations of mathematics.InPhilosophy of Mathematics , 2nd edn., ed.P.Benacerraf and H.Putnam.Cambridge: Cambridge University Press (1983) , pp.52-61.

Hibert, D. （1899）.Die grundlagen der geometrie.InFestschrift zur Feier der Enthullung des Gauss-Weber Denkmals in Göttingen .Leipzig: Teubner.

Horgan, T. （1994）.Transvaluationism: a Dionysian approach to vagueness.The Southern Journal of Philosophy , Supplement, 33: 97-126.

Hutton, C. （1795-1796）.A Mathematical and Philosophical Dictionary .2 vols.London: J.Johnson, and G.G.and J.Robinson.Reprinted, Hildesheim and New York: G.Olms Verlag（1973）, and in 4 vols., Bristol: Thoemmes Press（2000）.

Kant, I. （1781）.Kritikder Reinen Vernunft , ed.R.Schmidt.Hamburg: Felix Meiner Verlag（1990）.

Kitcher, P. （1989）.Explanatory unification and the causal structure of the world.InScientific Explanation , ed.P.Kitcher and W.Salmon.Minneapolis, MI: University of Minnesota Press, pp.410-505.

Klein, F. （1888）.Lectures on the Ikosahedron and the Solution of Equations of the Fifth Degree .London: Trübner&Co.

Klein, F. （1922）.The mathematical theory of the top（1896/97）.InFelix Klein Gesmmelte Mathematische Abhandlungen , ed.R.Fricke and H.Vermeil.Berlin: Springer.

Kreisel, G. （1967）.Informal rigour and completeness proofs.InProblems in the Philosophy of Mathematics , ed.I.Lakatos.Amsterdam: North-Holland, pp.138-186.

Laguerre, E.N. （1898）.Sur le calcul des systemes linéaires.InOeuvres de Laguerre , ed.C.Hermite, H.Poincaré and R.Eugene.Paris: Gauthier-Villars.

Lakatos, I. (1976).Proofs and Refutations , ed.J.Worrall and E.Zahar.Cambridge: Cambridge University Press.

Leibniz, G.W.F.Die philosophischen Schriften von Gottfried Wilhelm Leibniz , ed.C.J.Gerhardt and C.I.Gerhardt.Hildesheim: G.Olms Verlag (1978).

Leibniz, G.W.F.Discourse on Metaphysics and Other Essays , ed.D.Garber and R.Ariew.Indianapolis, IL: Hackett (1989).

Leibniz, G.W.F. (1764).New Essays Concerning Human Understanding , trans, A.G.Langley.Chicago: Open Court (1916).

Leibniz, G.W.F. (1683).Of Universal Analysis and Synthesis; or, of the Art of Discovery and of Judgement.InPhilosophical Writings [of]Leibniz , ed.and trans.M.Morris and G.H.R.Parkinson.London: J.M.Dent and Sons (1973), pp.10-17.

Leibniz, G.W.F.Opera philosophica quae extant latina, gallica, germanica omnia , ed.J.E.Erdmann.Aalen: Scientia (1959).

Leibniz, G.W.F.Opuscules et fragments inédits de Leibniz: extraits des manuscrits de la bibliothèque royale de Hanover .Paris (1903).

Leslie, J. (1809).Elements of Geometry, Geometrical Analysis, and Plane Trigonometry: with an appendix, notes and illustrations .Edinburgh: Brown and Crombie.

Leslie, J. (1821).Geometrical Analysis, and Geometry of CurvedLines: being volume second of a course of mathematics, and designed as an introduction to the study of natural philosophy .Edinburgh: W.&C.Tait, and London: Longman, Hurst, Rees, Orme, &Brown.

Lipton, P. (1991).Inference to the Best Explanation , 2nd

edn. New York: Routledge Publishing Company (2004).

Maddy, P. (2007). Second Philosophy: A Naturalistic Method. Oxford: Oxford University Press.

Mates, B. (1986). The Philosophy of Leibniz: Metaphysics and Language. Oxford: Oxford University Press.

Nelson, E. (1986). Predicative Arithmetic. Princeton, NJ: Princeton University Press.

Penrose, R. (1989). The Emperor's New Mind. Oxford: Oxford University Press.

Penrose, R. (1994). Shadows of the Mind: An Approach to the Missing Science of Consciousness. Oxford: Oxford University Press.

Penrose, R. (1997). On understanding understanding. International Studies in the Philosophy of Science, 11: 7–20.

Penrose, R, (2004). The Road to Reality: A Complete Guide to the Laws of the Universe. London: Jonathan Cape and New York: Alfred Knopf (2005).

Penrose, R. (2005). The Road to Reality: A Complete Guide to the Laws of the Universe. New York: Alfred Knopf.

Penrose, R. (2011). Gödel the mind, and the laws of physics. In Kurt Gödel and the Foundations of Mathematics: Horizons of Truth, ed. M. Baaz, C. H. Papadimitriou, D. S. Scott, H. Putnam and C. L. Harper, Jr. Cambridge: Cambridge University Press, forthcoming.

Playfair, J. (1778). On the arithmetic of impossible quantities. Philosophical Transactions of the Royal Society of

London , 68: 318-343.

Polkinghorne, J, C. (1996) .Beyond Science .Cambridge: Cambridge University Press.

Polkinghorne, J.C. (1998) .Belief in God in an Age of Science .New Haven, CT: Yale University Press.

Polkinghorne, J.C. (2005) .Exploring Reality .London: SPCK and New Haven, CT: Yale University Press.

Proclus.A Commentary on the First Book of Euclid's Elements , trans.G.R.Morrow.Princeton, NJ: Princeton University Press (1970) .

Putnam, H. (1967) .Mathematics without foundations.Journal of Philosophy , 64.

Putnam, H. (1975) .What is mathematical truth?InMathematics, Matterand Method , 2nd edn.Vol.1 ofPhilosophical Papers .Cambridge: Cambridge University Press (1979) , pp.60-78.

Quine, W.V. (1960) .Word and Object .Cambridge, MA: The MIT Press.

Resnik, M. (1980) .Frege and the Philosophy of Mathematics .Ithaca, NY: Cornell University Press.

Resnik, M. (1997) .Mathematics as a Science of Patterns .Oxford: Oxford University Press.

Rosen, G. (1994) .Objectivity and modem idealism.InPhilosophy in Mind , ed.J.O'Leary-Hawthorne and M.Michael.Dordrecht: Kluwer.

Rosen, G. (2006) .Review of Jody Azzouni, deflating existential consequence.Journal of Philosophy , 103: 6.

Rosen, G. (2010).Metaphysical dependence: reduction and grounding. In Modality: Metaphysics, Logic and Epistemology , ed.B.Hale and A.Hoffmann.Oxford: Oxford University Press.

Rosen, G.and Burgess, J.P. (2005).Nominalism reconsidered. InOxford Handbook of Philosophy of Mathematics and Logic , ed.S.Shapiro.Oxford: Oxford University Press.

Russell, B. (1905).On denoting.Mind , 14: 56.

Salmon, W. (1990).Four Decades of Scientifc Explanation .Minneapolis, MI: University of Minnesota Press, and Pittsburgh, PA: Pittsburgh University Press (2006).

Schopenhauer, A.Arthur Schopenhauer's Sämtliche Werke , Vol.2.Munich: R.Piper&Co.Verlag (1911).

Schopenhauer, A. (1859).The World as Will and Representation (Die Welt als Wille und Vortstellung) .New York: Dover Publications (1966).

Shapiro, S. (1997).Philosophy of Mathematics: Structure and Ontology .Oxford: Oxford University Press.

Shapiro, S. (2000).The status of logic. InNew Essayson the A Priori , ed.P.Boghossian and C.Peacocke, Oxford: Oxford University Press, pp.333-366; reprinted (in part) as 'Quine on Logic', inLogica Year book 1999 , ed.T.Childers, Prague: Czech Academy Publishing House, pp.11-21.

Shapiro, S. (2007).The objectivity of mathematics.Synthese , 156: 337-381.

Shapiro, S. (2007a).Vagueness in Context .Oxford: Oxford University Press.

Shapiro, S. (2009).We hold these truths to be self-evident:

but what do we mean by that?Review of Symbolic Logic , 2: 175-207.

Shapiro, S., ed.（2005）.The Oxford Handbook of Philosophy of Mathematics and Logic .Oxford: Oxford University Press.

Steiner, M.（1978）, Mathematical explanation and scientific knowledge.Nous , 12: 17–28.

Steiner, M.（1980）.Mathematical explanation.Philosophical Studies , 34: 135–152.

Steiner, M.（1998）.The Applicability of Mathematics as a Philosophical Problem .Cambridge, MA: Harvard Uiniversity Press.

Steiner, M.（2000）.Penrose and Platonism.InThe Growth of Mathematical Knowledge , ed.E.Grosholz and H.Breger.Dordrecht and Boston: Kluwer.Sternberg, S.（1994）, Group Theory and Physics .Cambridge: Cambridge University Press.

Tait, P.G.（1866）.Sir William Rowan Hamilton.North British Review , 14: 37–74.

Waismann, F.（1979）.Ludwig Wittgenstein and the Vienna Circle .London: Blackwell.

Waismann, F.（1982）.Lectures on the Philosophy of Mathematics .Amsterdam: Rodopi.-

Weyl, H.（1921）.ber die neue grundlagenkrise der mathematik.Mathematis-che Zeitschrift , 10: 39–79.

Wittgenstein, L.（1953）.Philosophical Investigations , trans.G.E.M.Anscombe, 3rd edn.Oxford: Blackwell（2001）.

Wittgenstein, L.（1956）.Remarks on the Foundations of

Mathematics , 3rd edn.Oxford: Blackwell（1978）; rev.edn., ed.G.H.von Wright, R.Rhees and G.E.M.Anscombe, Cambridge, MA: The MIT Press（1967）.

Woodward, J.（2009）.Scientific explanation.StanfordInternet Encyclopedia of Philosophy , http://plato.stanford.edu/entries/scientific-explanation.

Wright, C.（1992）.Truth and Objectivity .Cambridge, MA: Harvard University Press.

Lipton, Peter彼得·利普顿55

[1] 又译波尔金霍恩，珀金霍恩。——译注

[2] KBE是Knight Commander of the Most Excellent Order of the British Empire的缩写，英国高级骑士勋位英帝国勋位的一种爵士。——译注

[3] ANR是法语l'Agence Nationale de la Recherche（The French National Research Agency）的缩写。——译注

[4] 有中译本，《素数的音乐》，孙维昆译，湖南科学技术出版社，2007年第1版。——译注

[5] 有中译本，《神奇的数学》，程玺译，人民邮电出版社，2013年第1版。——译注

[6] OM是Order of Merit（功绩勋位）的缩写，英国高级骑士勋位之一。这种功绩勋章由爱德华七世于1902年设立，用于表彰英联邦健在的24名在国防、科学、艺术和文化领域做出杰出贡献的人士和1名外国人士。——译注

[7] 有中译本，《皇帝新脑》，许明贤，吴忠超译，湖南科学技术出版社，1996年第1版。——译注

[8] 有中译本，《时空本性》，杜欣欣，吴忠超译，湖南科学技术出版社，1996年第1版。——译注

[9] 有中译本，《通向实在之路》，王文浩译，湖南科学技术出版社，2008年第1版。——译注

[10] 有中译本，《数学哲学：对数学的思考》，郝兆宽，杨睿之译，复旦大学出版社，2009年第1版。——译注

[11] Alexander Grothendieck（1928—），法国数学家，生于德国柏林。他创立的概型理论将代数几何的研究提升到了一个全新的境界，被誉为20世纪代数几何领域最重要的进展之一。1966年荣获菲尔兹奖。——译注

[12] Paul Joseph Cohen（1934—2007），美国斯坦福大学数学教授，因提出称为“力迫法”的数学证明方法而闻名。他用这种方法证明了，连续统假说和选择公理都不能用集合论的策梅洛—弗伦克尔（Zermelo-Fraenkel）公理标准来证明。——译注

[13] 这一表述称为阿贝尔定理。——译注

[14] 格里斯更愿意称它为“友善的巨头”。——译注

[15] 阿罕布拉宫（La Alhambra），位于西班牙格拉纳达市东南的萨比卡山上。整座宫殿具有浓厚的摩尔（Moor）文化特征，是西班牙摩尔艺术的瑰宝，也是世界上最壮观的宫殿建筑之一，已被列为世界文化遗产。——译注

[16] A Mathematician's Apology，有中译本，译名为《一个数学家的辩白》，李文林、戴宗铎、高嵘编译，大连理工大学出版社，2009年第1版。还有中文评注本，《一个数学家的自白》，李泳评注，湖南科学技术出版社，2007年第1版。——译注

[17] “狂飙突进”是指德国（实际上不限于德国，至少也包括奥地利）在18世纪70～80年代发生于文学艺术领域的一场创作风格剧变的运动。在这一时期，德国的文学和艺术创作很快从古典主义过渡到浪漫主义。在文学创作上，强调个性、情感和进取性，代表人物有歌德和席勒，作品以歌德的《少年维特之烦恼》为先锋；在音乐上，强调突破形式（曲式），强调音乐的色彩和主观感受，海顿晚年的五首钢琴奏鸣曲，莫扎特、贝多芬和舒伯特的音乐创作都可以视为这一影响的先声。——译注

[18] The Library of Babel，一部由阿根廷作家博尔赫斯（Jorge Luis Borges，1899—1986）于1941年创作的一篇短篇荒诞小说。讲的是有这么一个由巨大的图书馆构成的宇宙。这个宇宙（或叫图书馆）由无数互相连着的六角形房间构成，房间内四壁都是书架，书架上垒满了由每本410页，由25个字符排列组合写出的绝无两本雷同的天书。曾经的图书管理员号称这些书包含了宇宙的全部知识，可现在就是没人能懂。本文作者杜·索托伊将篇名取为“探索巴别数学图书馆”也有这层意思。——译注

[19] 这个图首次出现在Penrose（1994），但我经常在其他地方用它，如Penrose（2004）。

[20] 应明确指出的是，哥德尔和科恩的结果，即证明连续统假设不依赖于标准的集合论公理系统这一断言，其本身并不回答这个假设在某种绝对意义上是否正确这样的问题，见Cohen（1966）。

[21] 关于这些问题讨论的更多细节，参见Penrose（2011）。

[22] 这并不是说在数学实践上不存在可能需要我们假定数学对象的存在性的方面。事实上，如果我们回到数学的适用性问题，那么我们可能需要考虑科学的确认是否成立，这不仅是说我们有关数学对象的假设的结果要有客观事实依据，而且这些假设的某些方面事实上确实为真。

[23] 据称格奥尔格·克赖泽尔认为，实在论在数学上的问题归结为“数学的客观性问题，而非数学对象的存在性问题”（Putnam，1975：1979，第70页）。如果数学客观性的这种分析是正确的，那么这两个问题终归是不能分离的。

[24] Noēsis，这是个希腊词，意指“才智”“理解”，广义地也可译成“思想”，与后文的aisthēsis（“感觉”）相对。本章中很多地方的“理解”均是指这个词意义上的理解。——译注

[25] 这里著录的中译本是李秋零译本（中国人民大学出版社，2011年7月第一版），以便读者查阅。原书著录（书后参考文献索引）为Kant，1781：1990，A50—51/B74—75。另，后文中引用的康德原文的译文均依据李秋零译本，特此致谢。——译注

[26] 类似论述见康德《纯粹理性批判》，第18页（Kant，1781：1990，Bxxvi）。

[27]　正如康德所说：……我能够思维我想思维的任何东西，只要我不与自己本身相矛盾，也就是说，要我的概念是一个可能的思想，即使我不能担保在所有可能性的总和中是否也有一个客体与它相对应。是，要赋予这样一个概念以客观有效性（实在的可能性），要求某种更多的东西。但这种更多的东西恰不需要在理论的知识来源中寻找，它也可能存在于实践的知识来源中。康德《纯粹理性批判》，第19页（Kant，1781：1990，Bxxvi，自注）。

[28]　但哥德尔明确否认我们的这种对数学真理的类知觉经验是基于感性知觉。称它为类知觉，只是因为它以一种接近于感官知觉作用于我们的方式被给予或强加于我们。

[29]　所谓数学概念（至少其中的一些）“独立于”我们的心理行为和倾向，哥德尔的意思似乎是（1）数学概念存在，并且即使没有我们的心理行为和倾向的作用也会继续存在；（2）我们的心理行为或倾向的任何变化不会自动导致这些概念的性质的改变。

[30]　哥德尔的确切表述是“（集合论的）这些公理强迫我们视其为真。”“哪些”公理？是哥德尔假设的存在一种唯一的，或可能是唯一最好的，集合论的公理化？我想未必。假定有一集合论语言L，那么下述两个断言——认为（1）存在可用L公式化的命题的可识别的集合Π，这些命题的明显性不由可用L公式化的其他命题导出，和（2）不是Π的所有元素都需要甚至应该被视为集合论公理化的公理——不必一定是不一致的。例如，有可能存在这样一种情形：Π的元素之间逻辑上有重叠，这使得将所有这些元素看成是集合论公理化的公理变得没有必要，甚至是不良的。关于集合或集合的不同概念还有可能存在其他思路，譬如采用将Π的特定元素与特定概念相联系的方法来划分Π的元素，等等。

[31] 当然，比Q1或Q2更基本的问题是如何理解感觉命题概念这样一个困难问题。我不打算在这里谈这个问题，因为我的重点是关于Q1和Q2及其与强迫现象的关系问题。

[32] 回顾可知，哥德尔曾坚持认为，“数学描述了一种非感性的实在，其存在独立于人类心灵的行为和倾向，只可以被感知”（Gödel，1951：1995，第323页，楷体强调为本文作者所加）。

[33] 我认为，我们应该认真对待熟悉/养成因素对数学上信念选择的可能作用。这里虽不适合对这个问题做深入展开的讨论，但它与数学理论化过程中广泛倡导的各种约束有共鸣，其中最重要的是所谓不变性原理，一条在19世纪早期和中叶由皮科克（Peacock,　G.）和其他人制定并大力捍卫的原理，一条为近世数学家广泛采用的原理。不变性原理的最新表述（或者说其一种变体）是由柯朗（Courant, R.）和罗宾（Robbins, H.）给出的：……在一个扩充的数域中的运算，其逻辑和哲学基础本质上是形式主义的；这种扩充必须通过定义来创建，这些定义是随意的。但是，如果不能在更大的范围内保持在原有范围内通行的规则和性质，那么它是无用的。R.柯朗和H.罗宾，《什么是数学》，中文版，左平，张饴慈译，复旦大学出版社，2005年第1版。第103页关于不变性原理的更多的讨论，见Detlefsen（2005，第273～278页）。

[34]　在结论一节里我会讨论另外两人——波尔查诺和戴德金——的观点。他们也都强调甚至是普通数学信念背后的由精心陶冶所带来的影响。

[35] 人们还普遍认为，前证明性研究（pre-demonstrative investigation），从推进证明研究使之变得更容易、更有效这个意义上说，应当是证明研究的准备阶段。例如，阿基米德（公元前287～公元前212）就曾向埃拉托色尼（Eratosthenes）推荐过用“力学”方法作为进一步证明研究的手段：……我认为有必要为你写下并详细解释……一种方法，它提供了一种使你能够用力学方法来开始对某些数学问题进行研究……这种方法甚至对定理本身的证明也是有用的。在我看来，对于某些事情，首先用力学方法来考察会看得很清楚，虽然后来它们是用几何方法证明的，因为用这种方法来研究并不等于实际的证明。但如果事情所寻求的某些知识已通过这种方法获得，那么用这种方法来提供证明就要比事先不知道这些知识而进行证明来得较容易。阿基米德（Archimedes，1993，第221～222页）

[36] 就是那种能够支撑起真正科学知识的证明。

[37] 命题Ⅰ～Ⅳ是：Ⅰ：用给定的有限长直线构造一个等边三角形。Ⅱ：过给定点（作为端点）作一线段使之等于已知线段。Ⅲ：给定两条不等长的线段，从较长的一条线段上截取一线段，使之与较短的线段等长。Ⅳ：如果两个三角形的两条边对应相等，且由这两条边所夹的角也对应相等，那么这两个三角形的底边也相等，两个三角形彼此全等，余下各角也对应相等。

[38] 命题Ⅳ也是关于全等三角形、角和等角等性质的定理。正因为如此，人们自然希望看到确立这些术语的存在性的基本命题。普罗克洛斯对此没有评述，对此我们只能怀疑他是否把这一点看作是欧几里得几何的缺陷了。

[39] 更具体地说，命题Ⅱ和Ⅲ为产生等长线段提供了两种基本方法：一种方法用于在给定位置上此前不存在线段的情形，另一种方法用于在那个位置上已存在较长线段的情形（见Proclus，1970，第183页）。

[40] 类似的观念在西方法学里，特别是在一般采用的犯罪事实原则中很常见。它要求对于合法定罪和正当审判都需要证据。这个观念是：一项审判如果是正当的，就必须在犯罪事实“存在”和犯罪者身份确认两方面都有证据。在谋杀案例中，这种证据通常以作为行为结果的死亡的证据形式出现，而确定犯罪嫌疑人的证据则是其实施的手段。其他类型的犯罪要求提供其他类型的证据。而所有合法的审判，都需要在犯罪事实存在和犯罪嫌疑人身份确认两方面有充分证据。

[41] 括号内数字为莱布尼茨著作的拉丁文本（1978，第7卷）的页码，类似陈述见Leibniz（1989，ⅩⅩⅣ）。

[42] 关于这些令人困扰的问题的简明而又用的讨论，见Mates（1986，第66～68页）。

[43] 这里当然是假设了A仍是一个真正的概念，尽管它是不可能的。当然，有许多人对此提出质疑。关于这一点后面会有更多的讨论。

[44] 莱布尼茨关于真实定义的例子就是欧几里得关于圆的定义——即“平面上一直线段绕一固定端运动所画出的图”（出处如上）。这在17、18、19世纪一直作为真实定义的一般性理解。例

如，赫顿在他的《数学和哲学辞典》里就提供了这种定义的如下特性：真实定义……解释一件事的起源，就是说，一件事是如何做或如何发生的：它就像圆的定义，我们将圆定义成由一直线段绕一定点运动所画出的图。赫顿（Hutton，1795～1796：2000，第1卷）

[45] 这里概念的“一致性”是指被用来描述或应用这一概念的陈述的一致性。

[46] 我认为，目前还不清楚，按照莱布尼茨和弗雷格的真实定义给出的构造性或具体例证的证明，在多大程度上类似于希尔伯特的模型构建证明。希尔伯特为几何提供的模型不是一种通过可视化或经验来说服人的模型。它远比欧几里得的圆的真实定义更抽象，它预设了一种做出抽象判断的能力。希尔伯特所运用的模型构建实际上是由直觉得来的，或者说在特性上推论性更强（more discursive），因此其本身需要直观的辩护这一点似乎很容易受到怀疑。

[47] 德文版见Schopenhauer（1911，第76页）。本书第1版出版于1819年，第2版1844年，第3版1859年。这里引用的段落见所有这些版本。

[48] 当然，在有关模型构造与代表真实定义的实例所展示的东西之间的相似性方面，存在严重的问题，但限于文章篇幅，我不能在此详作讨论。

[49] 戴德金在算术中提出过类似的关注。在《数是什么且应该是什么？》（*Was sind und was sollen die Zahlen*？）一书的第一段里，他因此对下述内在直觉的引导作用提出过警告，他说，看上去显然的真理往往是最需要证明的东西。这一思想的确成为他的算术基础研究的主线，其核心原则是：任何东西都应经过证明才能被接受，无论其显然的程度如何。见Dedekind（1888），第1版前言。

[50] 在这句话里，波尔查诺所说的实在是指经验的或可感知的实在。由此，他将欧几里得几何里的真实定义类型从科学的数学降低到当时被认为属于应用数学的范畴。

[51] 当然，罗素等人明确将这一点看成是集合论公理的“回归”概念。

[52] assertibility，这个词由弗雷格的术语assertion（断定）扩展而来。在弗雷格那里，“断定”是对一个语句的内涵的真假做出判断，而不仅仅是语法上的肯定句。由assertion还引申出诸如“assertion-sign”（断定记号）、“assertoric force”（断定力）之类的其他术语。有兴趣的读者可参阅有关弗雷格哲学的文献。——译注

[53] 这里的“世界”即上文所提到的小号大写字母所表示的世界，为区别起见，用黑体标示之。——译注

[54] 原书名为Il*Saggiatore*（*The Assayer*）。在中文出版的诸多科学史著作和评述伽利略的文献里，这本小书不是很出名，鲜有介绍，故值得在此说上几句。这本小书的肇始是要回击奥拉齐奥·格拉西就彗星问题引发的责难，后者化名洛塔里奥·萨西著书攻击伽利略1610年以来用望远镜做出的天文观测发现，企图巩固天地对立的传统宇宙观。于是伽利略于1623年写了这本小册子予以回击。从科学史的角度看，这本书之所以重要，不在于它写的关于彗星的论述以及伽利

略对自己制作望远镜的过程的细述和辩护，而是它在字里行间明确阐述了伽利略关于世界图景的深刻的认识论思想及其物理学基本信念。可能由于其成书时间早于伽利略著名的《关于托勒密和哥白尼的两大世界体系的对话》（1632），书中的很多思想在《对话》里有了更成熟的表述，因此后世往往略过了这本书。具体论述不在此赘述，有兴趣的读者请查阅《伽利略传》（B.G.库兹涅佐夫著，陈太先、马世元译，商务印书馆，2001年版）和最新出版的《近代物理科学的形而上学基础》（E.A.伯特著，张卜天译，湖南科学技术出版社，2012年版）。——译注

[55] 译文引自《近代物理科学的形而上学基础》（E.A.伯特著，张卜天译，湖南科学技术出版社，2012年版），特此致谢。——译注

[56] 范畴（category）是指用以表述思想、语言和实在的最基本的概念。而后文的“概念（concept）”是指由范畴导出的概念。——译注

[57] 尤西弗罗对比（Euthyphro contrast）也称尤西弗罗难题（Euthyphro dilemma），源自柏拉图的对话体著作《尤西弗罗》。这部书以尤西弗罗与苏格拉底对话的形式阐述了这样一个问题：“虔诚的行为是因为得到上帝的赞许而获得了虔诚的性质呢，还是因为它虔诚而博得上帝的赞许？”书中并未给出问题的答案，故称为尤西弗罗难题。它的实质是要挑明一个人的道德行为准则与善恶之间是否存在或存在怎样的因果关系。——译注

[58] 后面我们再回头来谈模糊性问题。

[59] 我在上面指出过，按照赖特的理解，认知律令仅当言谈受到认识论上的限制——仅当不存在不可知的真理——时才是客观性的一项基准。这里我们假设，数学受到认识论上的限制。为了给出这样一个机会，我们还必须只谈理想化的数学家。这些非人类数学家肯定能够知晓这样一些真实的命题，对这种命题人类数学家之所以不知晓，仅仅是因为命题所要求的计算太长，难以在太阳落山之前完成。

[60] 回忆一下：“言谈表现为认知律令，当且仅当我们可以先验地判定，观点的区别只有根据‘不同的输入’才能得到令人满意的解释，也就是说，争议双方是在不同的信息（因此负有无知或失误……的罪责）的基础上讨论，或是在‘不合适的条件下’（导致注意力不集中或注意力旁置并由此造成推理错误，或看错数据，等等），或是在‘出错状况下’（例如，对数据评估有偏见……或过于教条，或在其他范畴上出错……）讨论。”（Wright，1992，第92页）

[61] Monster—barring，见拉卡托斯《证明与反驳》（中译本，方刚，兰钊译，复旦大学出版社，2007年第1版）。——译注

[62] 在本文中，“存在”一词可按“独立于物质世界的真实存在”来理解。——译注

[63] 有关这一思想的最新陈述，见Dorr（2008）。

[64] 我可能只是简单说明了我在用“exist”和其他表示存在的习语时均按这些词在日常英语中的

意思来理解。正如奎因强调的，没有充分的理由相信这些词就意味着数学上的一个对象或其他领域内的某件事情（Quine，1960，§27）。是人都明白数不同于桌子、夸克和心理想象。但如果所有这些东西都是一种存在——即世间就有这种东西——那么我们就没有理由假设它们是以不同的意义存在着。

[65] 某些具有哲学头脑的数学家曾对经典算术的某些部分（例如，非直谓的或非构造性的部分，见Nelson，1986）表示怀疑。我们不必对这些具体化的争论太过在意。构造性算术包括存在性断言，例如断言在15和20之间存在两个素数。

[66] “算术就像是数的自然史（矿物学）。但是谁像这样来谈论它呢？我们的整个思维都贯穿着这种观念。”（Wittgenstein，1956：1967，Ⅳ，第11页）

[67] 这种观点在当代哲学家中信奉者寥寥无几，虽然数学家们常发现它很合己意。见Curry（1951）。有关弗雷格对形式主义的著名的反对意见的陈述和评判，见Resnik（1980）。

[68] 包括支配后继函数的基本原理在内的公理如下：0是一个数，0不是任何数的后继数，每个数都有一个后继数，没有两个数有相同的后继数。加法和乘法的递归方程为：对于任意数x，有x+0=x，对于任意数x和y, x+y的后继数=x+y的后继数；对于任意数x，有x×0=0，对于任意数x和y, x×y的后继数=（x×y）+x。数学归纳公理：如果0是函数F，且如果对所有的x, F（x）蕴含F（x的后继数），那么每个数都是F。

[69] PAω的完备性与哥德尔著名的不完备定理是相容的。哥德尔定理表明，任何包含基本算术的相容性理论必然是完备的，前提是这个理论是递归可枚举的。粗略地说，所谓一个理论是递归可枚举的是指列举其定理存在有限的机械过程——计算机程序。从这个意义上说，PA_{ω} 不是递归可枚举的，因此哥德尔定理对它不适用。

[70] ……或它在形式算术语言中的等价替代句子。

[71] 为明确起见，现在我们想象一种涉及两种示意性断言的形式主义：对于真的算术陈述A，（1）“A”为真这个事实是基于“A”是PA_{ω} 的一条定理这一事实；（2）A事实是基于“A”为真。

[72] 这种“排除性质的”结构主义模型见Benacerraf（1965），Putnam（1967）和Hellman（1989）。另一种也被称为结构主义的不同观点，见Resnik（1997）和Shapiro（1997）。

[73] 这里预设了自然数由0开始。

[74] 就我的目的而言，事实等同于伯特兰·罗素（Russell，1905）最初所描述的那种被构造的命题。

[75] 本小节所用的材料在Rosen（2010）里有更充分的叙述。

[76] 值得注意的是，还原论者也可能提出一种更强的断言，即，对于每一个数学命题p, p所处的情形正是q所处的情形，其中q是一个与更基本的对象有关的命题。考虑一种标准的科学还原情形：x比y热的情形正好对应于构成x的粒子的平均动能大于构成y的粒子的平均动能的情形。于是我们可以说，有关温度的事实强烈还原到有关平均动能的事实。这种强还原关系总是要求必然对等的断言：如果p强烈还原到q，那么作为一种必然性，p为真当且仅当q为真。这足以表明，强还原关系不同于我们前面讨论的支撑关系。[p或q]可能基于[q]，但没有人会说，p或q所处的情形正是q所处的情形。在一般情况下，[p]可能建立在[q]的基础上，即使[p]并非强还原到[q]，但反过来未必成立。如果[p]强还原到[q]，那么[p]必定基于[q]。在Rosen（2009）中，这种断言称为支撑—还原链接。

[77]　Emergent，指较高层级系统出现的性质（突现性质）不可能还原为较低层级系统的性质的一种状态。——译注

[78]　通常假设唯心主义和建构主义反对十分牢固的实在论，因为它们预先假定了某些表观上独立于心灵的事实具有心灵依赖的属性。就目前的解释而言，这只对了一半：这些观点是实在论的替代品，因为它们承认某些事实依赖于更低层次上与其他对象有关的事实。这些低层次事实包括心灵，这一点很有意思，但却不是主要的。算术哲学上的形式主义和建构主义也是十分牢固的实在论的替代品，出于同样的原因，在这些情形下，基本事实绝不可能是心理性质的。

[79] 关于用于分析的还原论建议的综述，见Burgess and Rosen（1997，第B章）。

[80]　在普通数学里对于这一点我们有先例。在某些基础性研究的语境中，对自然数235、整数235、实数235.0和复数235+0i等进行区分是很重要的。但是在许多日常语境中，这些差异并不重要。如果有人在这样一种基础语境中援引235+657=892这个事实，那么他的言谈无异于擦枪走火，就是说他无法用他的话挑选出一个简单的事实，虽然这种走火可能是无害的。

[81] 在现代数学哲学中强调“客观性”超过强调“对象”的起源，见Georg Kreisel的著作。

[82]　顺便指出，根据奎因的哲学，这个问题是毫无意义的，因为根据他所定义的柏拉图主义，如果没有对数学对象的“量化”，我们甚至无法描述可观察的世界。原子的不可或缺性则不同——我们不必求助于原子就可以描述其表象，但我们需要用原子的概念来解释这些表象。正如我已故的老师西德尼·摩根贝瑟（Sidney Morgenbesser）曾说过的那样，奎因的不可或缺性是某种形式的康德主义，这方面奎因可能是继承自C.I.刘易斯，因此说“奎因—普特南不可或缺性论文”是一种误导。普特南是在后一种意义上使用不可或缺性概念的。

[83] 本段引文见《第一哲学沉思集》（中译本，徐陶译，中国社会科学出版社，2009年第1版）中“反驳和答辩·关于第五个沉思”一节，译文采用徐译，特此致谢。——译注

[84]　我认为笛卡尔说的不只是康德所说的某些数学真理是人为的那么简单。在康德看来，无论是7+5=12，还是关于三角形内角和的定理，都是人为的真理。只是在第一种情形下，我们学不到有关数学本质的任何东西，因为总和不是这个数就是那个数。而一个三角形的角度之和才具

有真正的信息。

[85] 见http://www.chessbase.com/newsroom2.asp?id=239。感谢希尔万·卡佩尔提供了这个链接。

[86] 《皇帝新脑》中译本1996年第1版，第110～111页。这里的译文是本书译者按照原文并参考原中译文重译的。——译注

[87] 我很感谢Sylvain Cappell和Stanley Ocken在改进这里的陈述所给予的帮助。Cappell教授提醒我注意英国的拓扑学家弗兰克·亚当斯的深刻的结果：只有一维、二维、四维和八维欧氏空间可以构成满足乘法法则且满足下述条件的矢量空间：（a）存在一个乘积性单位矢量（左乘和右乘）；（b）两个非零矢量的积非零。当n=8时，乘法不满足结合律。

[88] 克莱因并未明确陈述过空间转动可以用行列式为+1的3×3实正交矩阵来表示，因此我们还不能将SU（2）二到一地同态到SO（3）的发现归功于他。

[89] 我们建立了自共轭矩阵与空间三维向量之间的另一种对应关系[具体见Sternberg（1994，§1.2）或Goldstein（1980，第4章）]。设A是一个自共轭矩阵，M是SU（2）的一个元。然后用M共轭，即MAM^*，来产生对应于一个转动向量的自共轭矩阵。

[90] 参见Hadamard（1954，第128～189页）：“最令人惊讶的——我应当说令人困惑的——是，这种特性与当代物理学的奇异标志联系在一起。1913年，埃利·嘉当，法国数学界的顶尖数学家之一，想到一类与群论相联系的解析的几何变换。在那个时候，除了其审美特性外，人们看不出有什么理由需要对这些变换做特殊考虑。然而，大约15年后，物理学家从实验上看到电子的一些奇异现象，这些现象只能借助于嘉当1913年提出的这一思想来理解。”

[91] 观察上的这种局限性使得开普勒不可能观察到一个行星受到的来自另一个行星的扰动，他发表的行星运动“定律”好像行星都不受这种扰动。这些定律使牛顿导出了引力的平方反比定律，然后又用之于研究行星的扰动。氢谱中“失踪”的辐射线实际上也一直都在那儿（只是较微弱），如果它们（在当时）能被检测到，那么我们就会错过一个重大的发现。

[92] 其优美形态见Chung and Sternberg（1993）。

[93] 这个例子在彭罗斯的《通向实在之路》（中译本，2011年第1版第2次印刷，第146～147页）中有清楚的图示和叙述。对类似的拓扑变换感兴趣的读者可以读一读中国数学名家姜伯驹先生写的科普小书《绳圈的数学》（大连理工大学出版社，2011年第1版）。——译注

第一推动丛书: 综合系列

未来50年

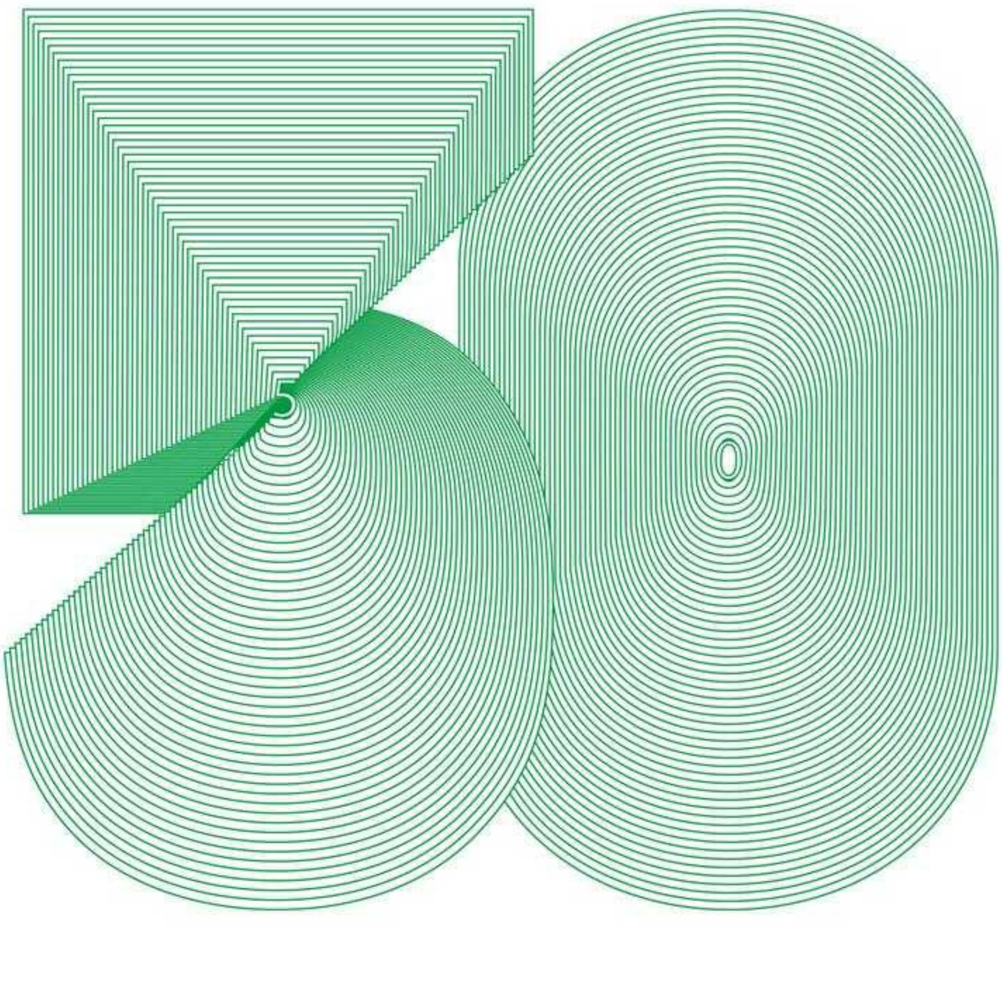
[美] 约翰·布罗克曼 编
李泳 译

The Polytechnique Series

The Next Fifty Years

John Brockman

THE FIRST MOVER



湖南科学技术出版社

版权信息

未来50年

编者:[美]约翰·布罗克曼

译者:李泳

责任编辑:吴炜　戴涛　陈刚　杨波

装帧设计:邵年　李叶　李星霖　赵宛青

ISBN：9787535794406

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
引言
第一部分 未来的理论
宇宙的未来
我们孤独吗我们在哪儿
2050年的数学
在文化的阴影里
替换大脑
儿童会告诉科学家什么
走向道德发展的理论
捉摸不定的心理学
幸福的未来
50年后我们还会忧伤吗
费米的“小发现”和复杂性理论的未来
生命是什么
第二部分 未来的实践
小摩尔定律
第二个创生
预测未来
肉体与机器的结合
物质的未来
我们会更聪明吗
复杂性的顶点
流动的信息
思想，大脑和自我
我们何以如此：来自2050年的观点
药、DNA和躺椅
心理疗法和穿在身上的电脑
征服疾病
译后记

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

引言

约翰·布罗克曼

纽约

2001年9月

1999年，我发表了一篇题为《第三文化》的文章，我在文章里提出一个观点：一个新的文化，一个大众的文化，已经出现了。它包括“那些科学家和经验世界的思想家，他们通过自己的工作和通俗的作品，生动表现了我们生活的更深层的意义，重新界定我们是谁、是什么。他们正在取代传统知识分子的地位”。

科学是大新闻，而提出大问题的人正是科学家。科学家通过他们的书和文章而成为新的大众的知识分子，成为新的大众文化的领头人。《未来50年》描绘了这个新文化的若干侧面。

展现在这里的文章，并不是站在边缘来讨论旧式的知识文化；这本科学家群体的作品集中讨论的是影响我们星球上每个人生活的那些发展。看看当今世界的那些出版物在说什么吧：干细胞研究、克隆、人类基因序列、人工智能、天体生物学、量子计算……这些话题（还有作品）必然都是跨学科的。在过去的十年里，越来越多的人在读科学家（包括本书的一些作者）写的书，原因之一是科学家被逼着用其他学科的同行能懂的语言来写东西。于是，受过普通教育的读者也能从中获益，因为在科学家担负起我们时代的大问题时，他也能参与进来，站在他们的肩膀上。

在这个文化里，在这本书里，科学家并不是为了迎合大众才写得普及的，他们那样写，是为了吸引我们时代的论战中的其他学科的同行们。他们的目标不是科学的普及，而是一种努力的尝试，为的是不仅让广大的读者理解最新的科学研究，而且在科学本身的意义下把它说得通

俗易懂。

换句话讲，对于我们日常生活里出现的问题，这些文章的作者并不一定能提出比普通人“更好的”答案。关键的区别在于他们所提问题的质量。

这25篇原创论文的主题和出发点是作者各自领域的“未来50年”。未来半个世纪的科学进步将如何改变我们的世界？如何改变关于我们是谁和是什么的问题？在每一个领域或学科，我们能期待哪些成果？它们又将如何影响和跨越其他学科？哪些期望是不现实的？哪些可能令人惊奇地改变我们的观念？

本书汇集了一篇篇有思想、有挑战的文章，一次次理性的历险。作者是25个一流的科学家，他们常常通过书籍和文章向大众普及他们的科学。他们是，生物学家Richard Dawkins, Paul W.Ewald, Brain Goodwin, Stuart Kauffman和Robert M.Sapolsky；化学家Peter Atkins；心理学家Paul Bloom, Mihaly Csikszentmihalyi, Nancy Etcoff, Alison Gopnik, Judith Rich Harris和Geoffrey Miller；心理学和计算机专家John H.Holland；心理学和人工智能专家Roger C.Schank；神经学家Samuel Barondes, Marc D.Hauser和Joseph Ledoux；计算机专家David Gelernter和Jaron Lanier；MIT人工智能实验室主任Rodney Brooks；数学家Ian Stewart和Steven Strogatz；天文学家Martin Rees；还有理论物理学家Paul Davies和Lee Smolin。

第一部分“从理论上”探索未来。内容包括：宇宙学进展，“虚拟非现实系统”在数学的应用，复杂性理论新方向，关于“活着”的意义，关于我们如何学习，如何思想，关于我们意识的本质，关于我们如何感觉，以及我们是否因为是宇宙中唯一的智慧生命而感到孤独。

第二部分“从实践”来看未来。内容包括：DNA排序的未来以及它所能告诉我们的关于我们自身的东西，火星探测与地外生命，我们对物质的控制，我们与机器特别是计算机的密切关系，虚拟空间的未来蓝图，神经科学，我们培养孩子的方式，我们不断进步的身体和精神的幸福前景。

我们正在穿越急剧变化的认识论的海洋。我们手握着威力空前巨大的新工具。正如牛津的生物学家杨（J.Z.Young）在1951年BBC电台的

雷斯演讲里说的，我们在这个过程中也变成了那些工具。我们现在还缺乏一种能像技术改变我们那样飞快改变它自身前提的知识文化。

《未来50年》是这个改变的开始的一部分，在这里，经验论与认识论在碰撞，所有的事情都变得不同——在这里，我们开始重新思考自身的本质，思考我们生活在怎样一个世界。那种协同的作用，存在于本书呈现的思想家们的工作中，也存在于他们为本书所写的文章里。

第一部分　未来的理论

宇宙的未来

L.斯莫林

Lee Smolin

我们被请来预言今后50年的科学状况。从过去几百年物理学和宇宙学前进的步伐来看，50年够长了；但要做一个到那时不显得那么愚蠢的预言，它似乎又太短了。回顾科学的历史，你会发现，人们提出的许多问题常常在后来的50年里就有了答案。不过，科学的进步一般总是缓慢的，我们今天大概还用着50年前的同行们用过的语言。

现在，让我们回过头来，看过去50年有些什么样的大问题。我个人看来，应该包括：

1）把原子核束缚在一起的强力的本质是什么？

2）决定放射性衰变的弱力的本质是什么？

3）宇宙的稳恒态模型对吗？或者，真的可能存在像盖莫夫（Gamow）和其他追随者们猜想的大爆炸吗？

4）质子和中子有内部结构吗？

5）为什么质子与中子存在很小的质量差，而电子比二者轻得多？为什么中微子没有质量？μ子是什么？谁让它来的？[1]

6）广义相对论与量子理论有什么联系？

7）理解量子理论的正确路线是什么？

我想我们可以自信地说，我们现在知道前4个问题的答案。我们还在为后面3个问题努力。不过，我们也没忘记已经回答了的问题；实际上，回答那些问题的方法形成了我们今天培养理论物理学家的基础。

然而，如果倒退100年，我们会发现，人们那时提出的许多问题，今天已经不再有人关心了。我不是一个够格的历史学家，不能列举物理学家们在20世纪之交提出的问题，不过他们很可能更关心以太的性质，而不太关心原子的性质。过了几年之后才会出现物理原子存在的证据——实际上，在1900年，很多物理学家并不相信原子的存在。另一些人，如马赫（Ernst Mach），认为这不属于物理学问题，因为原子是永远也观测不到的。至于天文学，1900年还没有证据说明存在远离我们银河系的星系，也没有人想过恒星为什么发光。所以，虽然20世纪50年代初的物理学家可能理解今天的物理学家的问题，但物理学家在50年代相互交流的语言，大概没有哪个20世纪初的人能听懂。

有时，科学在50年间的变化很小，所以需要我们预测在那以后将知道些什么。但是，也有些时期，科学飞速发展，那样的预测也就不需要了。在未来50年和100年间的某个地方，似乎存在一条地平线，超过那条线，关于科学进步的任何具体的猜想都将失去意义。

让我们歇息片刻，想想为什么会那样。部分原因可能是，50年大约等于一个科学家从开始研究到退休的时间，因而也是他们的科学生涯产生保守倾向的时期，这种倾向反过来会拖科学进步的后腿。科学是艰难的，我们科学家喜欢尽可能理解自己所做的事情；于是，除非迫不得已，我们总是愿意用我们已经认识了的思想和技术。另一个原因是，青年科学家的经历常常受临近退休的前辈的控制，那些老人在许多时候已不再活跃，因而不熟悉新的技术。精明的研究生不论多有想象力，也不敢大胆做本领域的权威老人们所不理解的事情。于是，为了思考我这个学科在50年后像什么样子，我想象我那些最聪明的研究生在他们的退休晚会上可能谈些什么。我猜想，假如没有什么解释不了的事实，他们掀不起20世纪初那样的革命，他们还会继续运用我们教给他们的语言。如果那样，现在的训练还是有用的——尽管我们中间的浪漫主义者更盼望一场革命，而不仅仅满足于对我们信仰的证明。

我们还可以想想，推动20世纪前50年科学取得那样巨大进步的社会背景有什么不同的地方。我想到了两个可信的答案：一个是，像爱因斯

坦和埃伦菲斯特（Paul Ehrenfest）那样的“门外汉”，尽管没有大学职位，也能发表他们的东西；另一个是，量子理论创立者的前辈们多数被第一次世界大战耽误了，为海森伯、狄拉克和他们的朋友留下了广阔的天地。

那么，未来50年里，关于基础物理学和宇宙学我们将知道些什么呢？我不做猜想了，而是提出一种方法，它可能得到在2050年看来不那么愚蠢的结果。我先列出今天还没回答的几个最基本的问题，然后谈谈能对这些问题的答案做出检验的实验和观测科学将有哪些可以预料的进步。我不担心理论的发展，因为我所列的所有问题都已经提出了理论的回答，我想在50年的时间里，我们理论家能够调整自己的理论，或者创立新的理论来满足实验和观测的数据。

下面就是我列出的7个最重要的基础物理学和宇宙学的问题。

1）现在形式的量子理论正确吗？它是否需要修正，要么达到一个合理的物理学解释，要么跟相对论统一起来？

2）引力的量子理论是什么？普朗克尺度下（10^{-33} 厘米，比原子核还小20个数量级）的空间和时间有什么结构？

3）决定基本粒子性质的那些参数的准确数值，包括它们的质量和相互作用的力的强度，由什么来解释？

4）用什么来解释我们看到的巨大数量级的比值？为什么两个质子间的引力比它们之间的电斥力小10^{40}？为什么宇宙那么大？为什么它至少比基本的普朗克尺度大60个数量级？为什么宇宙学常数比物理学中的其他参数几乎小同样多的数量级？

5）大爆炸是什么？从大爆炸中产生的宇宙的性质由什么决定？大爆炸是宇宙的起点吗？如果不是，在它之前发生过什么？

6）占宇宙密度80%到95%的暗物质和暗能量是由什么组成的？

7）星系是怎么形成的？我们观测的星系分布图像能告诉我们宇宙早期演化的什么情况？

前面4个问题从50年前就开始不断被提出、深化，没有结果。其余3个问题是新的。现在我们来看，我们在2050年进行的观测和实验是否足以检验理论家们为这些问题提出的答案。当然，在50年里什么事情都可能发生。不过，为了使我的方法更可信一些，我们只好在技术进步方面保守一点。所以，我只考虑已经存在和正在发展的技术。在后一种情形，我只考虑绝大多数专家认为可能在近些年产生作用的技术。不过，每一种技术，不论现有的还是在发展中的，我都假定它在未来50年里将只受物理学定律和经济条件的限制，而得到尽可能的发展。普通的显微镜存在光线波长的自然极限，而望远镜的自然极限来自有限年龄的宇宙中的有限的光速。其他技术可能更多受财政方面的限制。我们可以放心地假定，没有哪个实验（在那时）的消耗能超过美国的国防预算。不过我得赶紧声明，我不是实验物理学或观测宇宙学的专家，我没有认真研究过有关的技术极限。所以我的估计必然是夸大的。如果现有技术能发展到它们在自然和财政的极限状态，那么，下面就是我对未来50年的希望。

我们可以从量子理论说起。今天正发展着强大的新技术——主要是帮助发展量子计算机的技术，有可能极大地扩张我们过去实验检验量子理论的范围。它们是一些宏观机器，利用量子效应（如叠加和缠绕）来做普通计算机不可能做的计算。[2] 量子计算机要求那些目前只在原子系统观测过的量子效应为计算机线路那样的宏观系统服务，所以，那些机器能检验量子理论的那些最不同于经典理论的预言。

人们已经看到，量子计算机能破译现在政府、军队和商务使用的所有密码，所以金钱会源源不断地向量子计算机涌来。于是我们可以大胆地猜想，只要量子力学外推到宏观系统也是正确的，未来50年里会出现量子计算机，而且很可能出现利用在全世界无限延展的量子态的量子通信手段。另外，假如今天的量子理论只是某个更深的理论的近似，量子计算机的实验也可能证明这一点。因此，我们有理由相信，从现在开始的50年里，我们将知道第1个问题的答案。

我们接着来看宇宙学。到21世纪中叶的时候，通过全波段电磁光谱的观测，加上中微子、宇宙线和引力波，我们一定能有一幅详细的宇宙历史的图像。现在的宇宙模型里的参数将在很高的精度上得到测量，我们还将知道宇宙的许多其他事实，如黑洞的数量，恒星、星系、黑洞、中子星、类星体、γ射线源和其他天体在时间和空间的分布。实际上，

我们那时对宇宙的详细历史和性质可能比我们今天对我们行星表面的历史，知道得更多。至少，就全方位熟悉宇宙的现象来说，我们可能真的“认识到家”了。

结果将极大约束今天流行的关于早期宇宙的理论，如暴胀理论。关于星系如何形成，星系团和超星系团的模式如何形成，我们也将有一个具体的图像。即使不能直接观测暗物质，那些观测也将约束我们关于暗物质和暗能量本质的理论。到21世纪中叶的时候我们也许已经直接观测到了暗物质和暗能量，充分认识了它们，能证明或否定为它们提出的形形色色的理论；当然，我们也许什么也做不到。[3]

有读者会问，是不是所有那些观测都能证明大爆炸理论？为回答这个问题，我们需要区别大爆炸宇宙学的两种意义。我把第一种叫膨胀宇宙的宇宙学：这个理论说，宇宙从一个致密炽热的状态开始，已经膨胀了大约130亿年。在这一历史过程中，关键的一个事件是光与物质的分离，发生在宇宙冷却到原子能稳定下来的时候。在那之前大约100万年，宇宙像星体内部一样充满了等离子体。经过那个转变，宇宙充满了非常稀薄的气体，光可以通过，天变得透明，我们今天看见的所有结构——恒星、行星、星系和星系团——都形成了。而且，几乎所有的化学元素也都是从这个转变开始在恒星里形成的；只有氦和几种轻元素（如氘和锂）是在它们之前形成的。这样的图景，我想在未来50年里不大可能有什么修正。我们会更多地了解恒星、星系和元素的形成过程，但所有的证据仍然会支持这个膨胀宇宙的理论。

我们还可以满有把握地说，我们的观测将极大影响我们关于极早期宇宙历史的理论。如果让时钟倒转，宇宙的密度和温度会增大。有趣的问题是，回到什么时候我们才能通过观测来决定我们的理论。到21世纪中叶，面对检验的那部分理论可能至少要回到普朗克时间，那是一个极小的时间单位，10^{43} 个那样的间隔才等于1秒。我们来看一个暴胀假说的例子。在一系列合理的假定下，这个理论的预言可以通过今天宇宙微波背景的涨落的观测来检验。这些观测是现代科学的伟大贡献之一。不过，即使今天的观测能与暴胀理论相容，仍然存在许多没有解决的问题。暴胀的预言很简单，同样可以包含在其他理论中，所以，为了将暴胀跟今天事实的其他可能的有竞争力的解释区别开来，我们还需要更详细的测量。另外，暴胀理论有许多不同的形式，当然也需要更进一步的测量来区别它们。我们希望能在5年而不是50年里拥有宇宙微波背景的

更详细的观测。因此，我们有理由预言（当然也说不定），半个世纪以后，如果谁还想着可以通过回到遥远的普朗克时间进行观测来检验膨胀宇宙的理论，那就太落伍了。

但普朗克时间还不是时间的起点。跟膨胀宇宙理论迥然不同的还有另一种说法：大爆炸是宇宙的绝对开端。即使我们知道在某个理论起点过后若干分之一秒的遥远过去，宇宙更热也更密，也并不能证明在那之前不曾发生过推动宇宙膨胀的事情。所以，仍然有那样的可能：在理论的“大爆炸”瞬间之前，宇宙也许以不同的形式存在了很长的时间。为了鉴别这些不同的假说，我把它们称作宇宙起源的理论。

现在研究的宇宙起源理论有几个，都跟膨胀宇宙理论相容，从而也跟目前所有的观测结果相容。其中的一些，如哈特尔（Hartle）和霍金的“宇宙波函数”预言（更恰当说，是假定），认为大爆炸是时间的真正起点。另外的一些则认为新宇宙是在黑洞的坍缩中创生出来的，它们预言在大爆炸之前还有一个宇宙，而那时发生的事情决定了从大爆炸生出的宇宙的性质。大爆炸之前是否存在过什么，在这一点上我们也许能（但没有一点儿把握）拿出证据来约束那些理论。只有用引力波来探测最初膨胀时的宇宙，我们才可能获得那些证据。别的任何东西都不可能做到，因为初始的宇宙对任何形式的辐射都是不透明的，只有引力波例外。引力波天文学在发展着，但至今还没发现引力波（引力波由美国科学家在2016年宣布被LIGO探测到——编者注）。有许多案头的空间引力波探测器计划，原则上能用引力波为普朗克时间下的宇宙拍摄系列快照，从而区分不同的宇宙起源理论。到21世纪中叶时，这些技术也许能实现，但也说不定。

现在我们来谈基本粒子物理学。这里的限制来自经济。如果加速器技术没有突破，我们就不会制造比正在建造中的预算几十亿美元的加速器更加强大的机器。因为，加速器的能量是以对数方式随经费增加的，为了获得多10倍的能量，必须多花100倍的钱。因此，我们可以预言，到21世纪中叶时，除非有什么新的技术发明，我们能探测的能量最多比今天高两三个数量级（探测的空间尺度相应地小两三个数量级）。不过，即使这样不紧不慢地进步，也能带来很多发现，如希格斯（Higgs）玻色子——人们相信我们观测到的所有粒子都是通过它来获得质量的。我们还应该能证明或者否定超对称性假说，那是弦理论的基本要素。所有这些都很重要，但距离普朗克尺度还差15个数量级呢——

为了直接检验引力的量子理论，那是我们必须探测的尺度。那么，这是否意味着像弦理论和圈量子引力那样的候选量子引力理论在未来50年仍然不能落实呢？[4]

也许不是的！新技术已经有可能探测普朗克尺度了。原来是这样的：有些引力的量子理论预言空间和时间具有离散的原子式的结构。假如真是那样，光子通过空间的路线会发生改变——就像光线通过水时会发生散射和折射。这个效应极微弱，但能累加起来：假如光子经过很长的距离，小效应可以放大。幸运的是，我们很容易观测到那些来自高能事件（如γ射线爆发）的旅行了几十亿光年的光子。在这些和其他几种情形，我们完全可能观测某些量子引力理论预言的效应。实际上，已经有人提出，在甚高能宇宙线行为里观测到的某些效应，就可以用普朗克尺度下空间的量子时空结构的效应来解释。

应该强调一点，我这里说的观测，是为其他目标而设计的卫星所做的观测。在未来的10年，等我们有了为自己目标而设计的卫星，就可以探测普朗克尺度了。为了推进这些实验，我们需要提高探测器的精度，而据我所知，高能光子探测器的精度没有自然或经济的限制，因而新的方法可以无限地用来探测普朗克尺度的时空结构。就宇宙线而言，还是存在经济约束：随能量增大，宇宙线越来越少。所以，为了探测更高的能量，需要更大的探测器。然而，计划或建设中的探测器的研究范围，已经足以区分不同的量子引力理论；当然，好数据还是要等到2050年。

被认为最有可能成为引力的量子理论的候选者的弦理论，因为几乎没有做出可以检验的预言而遭到许多批评。然而它确实做出过几个预言，而且其中一个还跟这样的观测结果有关——那就是，结果必须与具有光滑结构的空间相一致。这是所谓洛伦兹不变性理论的对称性所要求的。其他的量子引力理论预言那种对称性是破缺的，或者被某些小效应改变了。这些不同理论的不同预言完全落在新实验能够检验的范围内。于是，过不了几年，弦理论或它的某个竞争者就可能被观测清除出局。

通过对实验潜力所做的匆匆考察，我们可以得到下面几点结果。我们可能会有很好的数据来决定前2个和后3个问题的答案，也就是关于量子理论和量子引力的问题，以及关于宇宙学和天体物理的问题。我们可能有，也可能没有检验宇宙起源理论的数据。引力波天文学可能揭示来自大爆炸之前的宇宙纪元（如果真有那样的一个时间）的信息；不过，

考虑到今天引力波天文学的摇摆不定，这一点还说不准。

那么，还剩下两个问题，第3个和第4个。它们问的是粒子物理学的参数，如基本粒子的质量和它们之间的相互作用的强度，为什么会有那样的数值。情况还不清楚。我们已经有了许多相关的数据，但现在还回答不了这些问题。下一代粒子加速器可能检验的那些粒子物理学进展，也许能帮助我们理解为什么基本粒子有那样的质量和相互作用。但是，多探测几个数量级是不是就足以解决那些问题了，还很难说。

也许出现另一种情形：这些问题都能用引力的量子理论来回答。这是弦理论的初衷，不过至今它还没有走上路来。目前的理论证据反倒说明，像弦理论那样的引力的量子理论似乎能与很多可能的基本粒子的性质相容。这是因为，我们寻求的答案所关心的是一个特殊理论的哪个解描绘了宇宙，而不是选择哪个可能的理论是正确的。每个理论似乎都有很多解，而每个解都描写一个可能的宇宙。

这就把有些人引向了存在多个宇宙的理论。或者说，只有一个宇宙，但包括了许多区域，每个区域都跟我们的宇宙相似。每个区域都从大爆炸开始，然后膨胀，生成物理学原理决定的结构。大体上说，有两种类型的多宇宙理论。第一种只假定世界由大量宇宙组成，每一个宇宙里的自然定律（或至少是它们的参数）都是随机产生的。我们一般把这种观点称作人存原理。第二种理论假定了一个过程，在那个过程中，新宇宙是作为黑洞形成的结果而产生的。这个所谓的宇宙自然选择的理论很像进化论的生物学，因为最普遍的宇宙是那些自我复制最多的宇宙。

在这些理论中，人存原理是完全不可能检验的，而自然选择假说有可能被证伪却不可能被证明——至少，在今天的技术条件下是这样。为了证明而不是否定那个理论，我们需要探测大爆炸之前的时间。引力波也许能为我们做到这一点，但正如我说过的，那可能是50年以后的事情了。所以，尽管自然选择的思想也许能经受2050年以前的一切考验，但它不会得到证明。

当然，新的思想和技术的出现也许会极大改变这种状况。不过我们已经说过要保守一点，只考虑现有的思想和技术。假如要我在保守分析的基础上大胆猜测，我想在50年后我们至少能回答上面提出的7个问题里的5个。至于我们能否回答问题3和问题4，任何人都可以有自己的猜

想。就是说，我们也许能知道引力的量子理论，认识大爆炸的本质，发现并认同量子理论的正确形式，但我们还是不能回答早在20世纪30年代就提出的那个简单问题：为什么质子和中子具有几乎相同的质量？为什么重一点儿的那个会是中子？

斯莫林（**Lee Smolin**）

斯莫林（Lee Smolin），理论物理学家，安大略沃特卢边缘研究所的创立者之一；《宇宙的生命与通向量子引力的三条途径》（The Life of the Cosmos and Three Roads to Quantum Gravity）[5]一书的作者。

我们孤独吗我们在哪儿

M.里斯

Martin Rees

未来50年最大的探索性挑战既不在物理学，也不在（地球的）生物学；它一定在寻找证明或者否定地外智能存在的确凿证据。我想，假如到2050年我们还不明白地球生命是怎么开始的，那就太奇怪了。那时候，即使没有来自地球以外的直接证据，我们也能估量其他行星曾经出现某些基本形式的生命有多大的可能。不过这紧跟着另一个看来更难对付的问题："假如简单生命出现了，那么是什么东西阻碍它进化成为我们认为有智慧的生命呢？"

在未来的10年里，我们要向火星发射一系列太空探测器，研究火星的表面，最后把样本带回地球。我们还有更长远的计划，用机器人探测器来研究太阳系的其他地方——例如土卫六（Titan）的大气和木卫二（Europa）冰下的海洋。假如有哪个仪器发现在我们这个行星系统的另一个地方独立出现过哪怕最简单的生命，也意味着那种简单生命一定更广泛地存在于银河系和更远的星系。

今天没人指望在太阳系的其他地方存在"高等的"生命，但我们的太

阳只不过是银河系几十亿颗恒星中孤零零的一颗。在其他恒星周围旋转的行星会不会栖息着比我们可能在火星发现的更有趣、更奇特的生命形式呢？那里是不是还生活着我们所谓的智慧生命呢？即使宇宙到处都有原始生命，高等生命也未必能出现。在这个问题上，我想不可知论大概是唯一理性的态度。我们对生命起源还知道得不够多——至于自然选择是不是“趋同”，地球重新演化会不会产生全然不同的结果，我们知道的更少——还说不清楚智能的“外星人”是可能还是不可能。

寻找来自地外智能生命的信号的努力，曾经很难得到公共的经费——甚至不如一部科幻影片的票房税收——这个话题总是讨厌地跟UFO等东西牵扯在一起。不过幸运的是，在一些爱幻想的人物的资助下，加利福尼亚SETI研究所[6]还在继续开展着他们的先驱性工作。

我对这些追寻很感兴趣，而且会一直坚持下去，尽管我相信成功的机会很小，因为其中混杂着显然的人为信号——甚至像一串素数那样单调乏味的东西。我们当然不知道那信号究竟来自“有意识的”生物，还是来自早在信号发出之前就已经灭绝了的什么物种留下的机器。但是这种接触将证明逻辑和物理的概念在人类大脑以外的其他地方也出现过。假如我们知道一颗遥远的恒星也像我们的太阳一样照耀着一群伴随它的行星，其中的一颗有着跟我们地球一样纷纭复杂的生物圈，我们大概会怀着新的情感遥望它吧。

这些追寻当然也许注定会失败。这在某些方面肯定是令人失望的。当然，失败并不意味着不存在其他智能生物——他们可能过着沉思冥想的生活，不愿做任何事情来表现他们的存在。另一方面，假如我们这个小小的地球真是智能生物的唯一居处，我们会大大增强在宇宙的自尊：跟银河系到处是复杂生命的情形相比，我们能坦然地以不那么卑微的宇宙观来看地球了。我们可以把地球看作“种子”，生命从它洒向整个空间。播种的时间是足够的，延伸10万光年的银河系，用不了我们从最早的灵长类生物走出来的时间，就能洋溢绿色的生机。太阳死的时候，将走过50亿年：那是地球从第一个多细胞有机物进化到今天的生物圈（也包括我们）所经历的时间的5倍。在那样漫长的世代里，也许还能出现更大的质的飞跃。未来的变化如果在人的指引下实际上会快得多，所以它们将发生在我们的文化和历史的时间尺度。我们不能预言生命最终会为自己塑造一个什么样的角色：它也许灭绝，也许兴旺起来，影响整个宇宙。后一种情形是科幻小说的话题，但是也不能因为荒唐而否定它。

从其他世界到其他宇宙？

假如我们回顾在20世纪所取得的成就，就不会觉得在未来50年确立上面说的某些猜想是什么可怕的事情。100年前，星星为什么闪烁都还是一个谜。那时，我们把银河当作一个静态的系统，根本没想到它的外面还会有什么东西。最近30年来，空间探测器发回了我们太阳系的所有行星的图片；通过巨大的望远镜，天文学家看到了比过去更加遥远的太空深处。我们的宇宙图景比我们能看见的最远的恒星还远几百万倍——来自最远的星系的光需要经过100亿年才能到达我们地球。宇宙的历史可以回溯到开始的1秒钟之前。只有在宇宙膨胀的第一个百万分之一秒以前，只有在黑洞的内部，我们才能遭遇未知的基本物理学。

这些进步引出另一种可能的迷人图景：我们所谓的宇宙——天文学家能探测的那个从大爆炸中产生出来的区域——不可能是全部的实在。根据确定然而纯思辨性的假设，理论家们已经为多重宇宙描绘了蓝图。林德（Andrei Linde）、韦伦金（Alex Vilenkin）和其他一些人通过计算机模拟描绘了一个“永恒的”暴胀时期，在那个时期里，许多宇宙从不同的独立的大爆炸产生出来，形成互不相交的时空区域。古斯（Alan Guth）和斯莫林则根据不同的观点，提出新的宇宙可以从黑洞内部产生，扩张成一个新的我们无法接近的空间和时间的区域。[7] 兰多（Lisa Randall）和桑德鲁（Roman Sundrum）猜想，其他的宇宙可以存在于跟我们分离的多余的空间维度里；那些互不相交的宇宙可能通过引力发生作用，否则它们之间就不会有任何相互的影响。用那个老式的类比说，如果气球的表面代表一个嵌在我们三维空间里的二维宇宙，那么其他的宇宙可以拿其他的气球表面来代表。在这样的一个表面上爬行的小虫，如果没有第三维的概念，不可能知道还有跟它一样的伙伴在别的气球表面爬行。其他的宇宙是跟我们分离的时间和空间的区域。我们甚至不能有任何意义地说它们是与我们同时存在，还是存在于我们之前或者之后；只有当我们能为所有的宇宙赋予同一个时间度量，那样的概念才会有意义。古斯和哈里森（Edward Harrison）甚至猜想，在实验室里让一堆物质坍缩成一个小黑洞，也能生出一个宇宙来。我们的宇宙会不会正是其他宇宙的实验产物呢？斯莫林猜想，主宰儿女宇宙的定律可能烙下了在父母宇宙中流行的定律的印迹。假如真是那样，虚构的神学主题将披着新的外衣而复活，从而进一步抹去自然与“超自然”现象之间的虚假的界线。

在“多世界”的理论中还有平行宇宙，是埃弗雷特（Haugh Everett）和惠勒（John Wheeler）在20世纪50年代为了解决量子力学的某些疑难而首先提出的。这样的概念，科幻小说先驱斯塔普里顿（Olaf Stapledon）早就描绘过了，那是他的造星者更精心的一个创造：[8]

> 每当一个生命面对几种可能的行动路线时，它会全部接受下来，从而创造出许多……不同的宇宙历史。因为在每一个宇宙的演化序列中都存在许多生命，而每一种生命都在不断面对着许多可能的路线，它们的所有路线的组合是数不清的，于是从每一个短暂的序列的每一个瞬间，落下来无限多个不同的宇宙。

这些图景没有一个是凭空幻想出来的，每一个都有严肃的（尽管假想的）理论根据。然而，它们最多只能有一个是对的。很可能没有一个是对的，还会存在别的恰好能带来一个宇宙的理论。

为了牢固树立起这些思想，我们还需要一个能和谐地描写宇宙开端的极端物理状态的理论——在那个开端的时刻，我们今天的宇宙还是在量子涨落中波荡地挤压在一起的小东西——换句话说，我们需要一个协调爱因斯坦的引力论（广义相对论）与微观亚原子世界的量子原理的理论。

爱因斯坦个人的最后30年都在追寻这样一个统一的理论，但是没有结果。刻薄的批评者认为他的后半生“最好去钓鱼”。现在我们知道他的追求太超前了，因为他还不知道核力和弱力，任何一个理论都必须把它们跟电磁作用统一起来。统一理论的挑战在21世纪很可能比它们在20世纪更加现实，现在似乎是向它们真正发起进攻的好时候了。但只是结合还不够，还必须有一个基础，好令人相信统一理论并不只是一个数学结构，而且还能用于我们身外的实在。假如我们所看到的事物——如核力、电力和引力间的联系，如为什么仅有三种中微子——都能用这个理论来解释，而且没有别的解释，那么我们对理论的信心也就树立起来了。在未来的几十年里，物理学家很可能建立一个这样的理论。那样的话，从牛顿以前开始，经过麦克斯韦、爱因斯坦和他们的后辈的那场理性的追寻，就走到尽头了。不过，假如它能发现宇宙为什么具有令我们某些人吃惊的特征，我会更加激动。

我们独特的热爱生命的宇宙

假如我们的宇宙里生活着异类的生命，他们也许会以跟我们截然不同的思维方式来理解和“整理”五花八门的实在。但是一旦建立起联络，我们一定会达成共同的兴趣。他们可能由同样的原子构成，遵从同样的物理学定律。假如他们也长着双眼，那里也有着明媚的天空，他们肯定会跟我们一样仰望星辰。我们为了自己的起源而追溯一个共同的创世纪——那个大约发生在130亿年前的大爆炸。

但是，不论我们的存在，还是异类生命（假如真有的话）的存在，都依赖于我们特别的与众不同的宇宙。一个适合生命的宇宙——我们也可以说一个热爱生命的宇宙——似乎一定是通过特别的方式组织起来的。任何形式的生命的先决条件，如稳定长寿的恒星、能结合成复杂分子的原子（例如碳、氧、硅）等，都强烈地依赖于物理学定律，依赖于宇宙的大小、组成和膨胀速率。假如大爆炸时留下的“配方”稍有不同，就不会生出今天的我们。许多配方的宇宙可能没有原子，没有化学，没有行星，还没出生就已经死了；还有些宇宙，可能太短命，太空虚，除了空空的一片，什么演化也不会发生。独特的配方似乎是一个根本的奥秘。

爱因斯坦关心的深沉问题是，“上帝在创造宇宙的时候有过什么选择吗？”他用诗的语言表达了这样的思想——统治我们宇宙的定律是否是唯一的（因为某个深层的数学理由），或者，原始的配方是否可以根本不同？假如我们的宇宙是一个基本理论的唯一结果，我们就必须接受一个铁的事实：宇宙是为了生命而特别组织起来的。另一方面，如果爱因斯坦的问题是肯定的，那么基本定律会有更多的可能：它们允许有很多的配方，产生很多不同类型的宇宙。整个的多重宇宙由一组基本原理来统治，但我们所说的自然定律（或至少它们的一部分）却只能是局部的律令，是我们那场特殊的大爆炸之后的最初瞬间的偶然的环境事件的结果。

我们可以拿雪花来做类比。所有雪花都有一点共同的地方：六边形对称。这种对称是构成雪花的水分子的形态产生的结果。但是很难找到两朵完全相同的雪花。每一朵雪花的花样都取决于它自己的历史。例如，在它通过形成它的云层的时候，周围的气温和压力以什么方式变化。同样，我们宇宙的某些特征也可能是大爆炸以后不同冷却方式的偶然结果，而不是整个多重宇宙留下的什么更深层、更基本的印迹。它像一块炽热的铁在冷却中磁化，而磁化的排列方式却依赖于一些随机因

子。如果物理学家得到了令人信服的基本理论，它应该告诉我们自然的哪些方面是那个基本理论的直接结果（正如雪花的对称花样是水分子基本结构的结果），哪些是偶然事件的产物（如每朵雪花的不同花样）。

假如确实存在一个宇宙的集合（能描绘我们大爆炸起点的理论也许能解决这个问题），那么大多数的宇宙要么是荒芜的——统治它们的定律排除了所有复杂的结构——要么太小或者太短暂，不允许时间和空间有充分的演化来产生复杂。我们（以及可能存在的异类生命）会发现我们生在一个不寻常的小世界，这里的定律允许复杂的演化。我们宇宙的这种看似“精心设计”的特征不应该令我们有多么惊奇，更惊奇的是我们在宇宙的特殊位置。我们不应跟着哥白尼的“平凡原理”走得太远。我们发现自己生活在一颗拥有大气的行星，在一个特殊的距离围绕着它的太阳旋转，凭这些，它就是一个非常特殊的与众不同的地方。在空间随机选一个位置都会远离任何恒星——很可能在星系间的某个虚空的空间，离最近的星系也有数百万光年。如果有很多宇宙，多数都不能居住，那么我们发现自己在一个可以居住的地方，一点儿也不应该奇怪。

总有一天我们会有一个令人信服的理论，它能告诉我们多重宇宙是否存在，某些所谓的自然定律是否只是我们这个宇宙小碎片的局部法则。即使在拥有那样的理论之前，我们也能通过下面的问题来检验这种“人择”的合理性：在可能产生我们的那个小宇宙集合中，我们这个现实的宇宙是不是典型的？假如即使在这样的集合里（不用说在多重宇宙的大集合里）它也是异乎寻常的，那么我们必须抛弃多重宇宙的假说。也可以从另一个方面来检验多重宇宙理论。我们考虑斯莫林的猜想：新宇宙诞生在黑洞的内部，儿女宇宙的物理学定律还保留着父母定律的记忆，仿佛有什么宇宙的遗传。如果斯莫林是对的，能产生更多黑洞的宇宙将具有更大的“生育优势”，而这优势还将遗传给下一代。如果我们的宇宙是这样一代代产生出来的，对生成黑洞来说，它应该接近最佳的状态。就是说，任何定律和常数只要有了丝毫的改动，黑洞就可能不那么好形成了。斯莫林的思想还没有得到具体理论的支持——如物理信息（甚至时间箭头）如何从一个宇宙传递到另一个。我个人对它的复兴也不抱多大希望。不过，我们还是应该感谢斯莫林，他向我们说明多重宇宙的理论在原则上是可能被否定的。

这些例子说明，某些关于多个宇宙的要求，跟任何一个好的科学假说一样，是可能被拒绝的。[9] 我们不能满怀信心地说发生过多个大爆

炸——我们连自己宇宙的极早期历史都没有足够的认识。我们也不知道基本原理是否是多种多样的，解决这个问题是物理学家未来50年的挑战。但是，假如基本定律是多样化的——假如它们允许各种宇宙的创生——那么，关于“我们的宇宙为什么那样”的所谓人择解释就合理了。实际上，对宇宙的某些重要特征，这将是我们所能得到的唯一一种解释。宇宙学的局部似乎会变成进化论的生物学。

我们传统意义的宇宙也许是许多大爆炸之一的产物，正如我们的太阳系不过是银河系里众多行星系统里的一个。池塘里冰晶的模式是偶然事件形成的，不是水的基本性质的结果；同样，某些看似不变的自然常数也可能是任意的，而不是基本定律唯一确定的。所以，为我们所谓的自然常数追寻精确的公式，可能终归是徒劳的，就像开普勒当年为行星轨道寻求什么精确的数字关联。[10] 别的宇宙也许成为科学进程的一部分，正如“别的世界”已经存在好多世纪了。不管怎么说（在这里，科学家应该高兴地把领地让给哲学家），关于为什么事物存在——为什么有宇宙而不是没有——的认识，还在形而上学的领地里，而且一定会永远驻足在那里。

里斯（**Sir Martin Rees**）

里斯（Sir Martin Rees）是剑桥国王学院的皇家学会教授。30岁时，继霍伊尔（Fred Hoyle）之后，在剑桥当选为天文学和实验哲学的Plumian教授。他提出过许多重要的宇宙学概念。例如，他第一个建议类星体奇异的高能量核心是从巨大黑洞获得能源的。在最近20年里，在剑桥的天文学院主持了一个宏大的研究项目。他写过几本书，包括《宇宙魔力》（Gravity’s Fatal Attraction，与Mitchell Begelman合作），《天体物理学新展望》（NewPerspectivesin Astrophysical Astronomy），《开始之前：我们的宇宙及其他》（Before the Beginning: Our Universe and Others），《6个数：形成宇宙的力》（Just Six Numbers: The Deep Forces That Shape the Universe），以及最近出版的《我们的宇宙家园》（Our Cosmic Habitat）。

2050年的数学

I.斯特瓦特

Ian Stewart

在所有的科学中，数学大概具有最悠远而连绵的历史——只有天文学能跟它比。两种学科至少都可以追溯到古巴比伦时代，那时的发现在今天依然是重要的。天文学建立在过去的发现上，数学也一样。天文学的基础是对现实世界的观测，而数学则是共有的思想的社会结构；但思想是天文学的驱动力，而数学却在对真实世界的模拟中成长起来——它记数过去的日子，测量田地的大小，计算给国王的贡品。

在天文学里发生过几次革命。旧的概念被推翻了，新的迥然不同的概念出现了。例如，1877年，意大利天文学家斯基帕雷利看到了火星的*canali*（“河道”），这个发现的误译很快传开，使很多人（甚至一些天文学家）相信火星上居住着智慧生命。[11] 现在我们有了更好的认识。人们常说，数学不可能有革命，因为数学真理的本质是不会改变的。但是，人的态度在改变，数学中最大的革命之一就是改变我们关于数学“真理”的概念。因为哥德尔（Kurt G del）和图灵（Alan Turing），我们才发现原来数学的真理也不是绝对的。

未来50年里，数学将发生几场大的革命。有的已经在发生了——计算机不断增大的影响，生命科学和金融行业提出的新的挑战。还会出现别的，但我们只能说，许多事情是不可能预言的。不同的评论家都预言数学证明的观念会发生改变，那可是数学的核心概念。有些人说计算机将带来一个根本不同的数学证明的概念，还有些人则认为那样的概念将彻底消失。两种观念都在根本上错看了当前的潮流。在数学中，证明取代了其他科学中观察和实验的地位——就是说，数学通过证明来避免被个人的聪明引向歧路，避免因为喜欢而相信并不真实的东西。显微镜的发明不能取代生物学实验，计算机同样也代替不了数学证明。正如我们在这个生物学的类比中看到的，计算机修正和强化了证明的技术，但是没有改变基础的哲学——证明说的是逻辑的一贯性，从已知的定理导出新的定理，而推导的路线应该能经受任何怀疑专家的最仔细的审查。在未来50年里，证明的概念仍将完整地保留下来，我们仍然相信它是数学事业中最基本的东西。

数学的力量来自两个不同源泉的汇流。一个是“真实的世界”。开普勒、伽利略、牛顿等告诉我们，外在世界的诸多方面可以通过简单而微妙的数学法则（“自然定律”）来认识。有时物理学家会修正这些定律的形式。牛顿力学让位给量子力学和广义相对论，量子力学让位给量子场论，量子引力和超弦指引着未来的理论修正的方向。现实世界的问题激发新数学的产生，即使产生它的理论改变了，那数学往往还在，而且依然重要。

数学的第二个力量源泉是人类的想象力——为了数学而追求数学。从真实世界到一个圆满的数学分支要走过一段艰辛的路程：需要经过一定的探索，而勇敢的先驱者们常常在追求个人的幻想中脱离主流，然后发现更好的路线。对这些先驱者来说，探索的价值是显而易见的：那正是他们的动力，除了它本身的意义，不需要更多的理由。

这两种作风的数学通常被划分为应用数学和纯粹数学。这两个词都不够准确，都是容易误会的概念。许多“应用的”数学实际上没有应用于任何现实的事物；“纯粹”数学的纯粹指的是它的方法，而不是说它轻视这门学科的实用价值。不过，这两个名词的确说明了不同数学风格的两个极端——而全部数学则是联系外在世界的规则与人类神奇想象力的一个统一的整体。正是这样的整体性和它双向的思维路线，才给数学带来了那么巨大的威力。我们要进步就需要那两种作风，硬说一个比另一个更优越是毫无意义的。

100年前，很多数学家拓展了那个数学体。才过50年，它已经庞大到没有人能完全把握了，于是个人越来越专业，生出四分五裂的学科。纯粹数学家与应用数学家分化为两大阵营，各自怀有不同的哲学。关于基础，关于证明的需要，关于方法，关于问题的意义，他们都抱着不同的看法。他们仿佛是从一个大的教派分裂出来的两个派别。但是在新千年到来时，这种自我分裂的趋势发生了逆转。纯粹数学的方法为应用数学带来了新的活力；应用中出现的问题刺激了纯粹数学的新发展。两家的界线开始模糊了，其实，那本来就不是什么实在的界线，而主要在于认识的分歧。在未来的50年里，涌向更大统一的潮流将越来越快，那时候，我们将只有数学家，没有定语的限制，没有派别的争论。专门化的专家还是存在的，不过，他们的专业将融合纯粹数学的抽象逻辑和概念意识与应用数学的具体考虑。我们都将成为数学家，为了那共同的伟大目标奋斗，徜徉在伟大的“超智慧”集合的数学的一块小小的自我天地中

间。我们将认识徜徉在其他小天地的伙伴，我们会感谢他们的存在，尊重他们的活动，因为他们的贡献，数学才更加光大。

关于未来50年，有一点我们可以满怀信心：我们会看到巨大的进步。数学的黄金年代不在古希腊，不在文艺复兴的意大利，也不在牛顿的英格兰，而在今天。在50年里仍然跟今天一样。这个观点的最好证明是在未解的大难题上取得的进步。那些问题提出几百年了，曾一直困惑着一些伟大的头脑，直到有一天我们发现了走近它的路线，产生了新的思想，那难题才张开缺口。最近，怀尔斯（*Andrew Wiles* ）证明了费马大定理，这是一个最好的例子。 [12] 大约1637年，费马（Pierre de Fermat）在他的丢番图（Diophantus）《算术》（Arithmetica）的页边写道，两个完全立方数之和不可能等于另一个完全立方数，四次方以及更高的次方，都是如此。几百年来所有关于这个问题的证明的努力都失败了。到1995年，怀尔斯才赢得了20世纪的这场最伟大的数学胜利。他的解决运用了一个新方法：将费马的表述转换为一种意义更加广泛的关于“椭圆曲线”——一个截然不同的数论领域——的命题，然后把每一种可能的现代工具都用进来对付这个新生的问题。

当前，最有名的一个未解难题是黎曼猜想，这首先是黎曼（Georg Bernhard Riemann）提出来的。这是复分析里的一个相当专门的问题，它猜想的答案可能为素数理论、代数数论、代数几何甚至动力学带来曙光。近些年来，还出现了它跟量子物理学的有趣联系。我想大胆预言，到2050年时，黎曼猜想会得到证明——人们期待的那个结论是正确的——而它与物理学的联系将在证明中发挥巨大的作用。不过，保守些说，我想解决猜想的最后路线不会基于它今天跟物理学的联系，真正的联系还难以想象。

1900年，那个时代最伟大的数学家希尔伯特提出了未来需要解决的23个重大问题。多数问题都解决了，但黎曼猜想还没有。2000年，麻省剑桥的克莱（Clay）数学研究所提出了7个久远的难题，每个问题悬赏100万美元。 [13] 其中一个就是黎曼猜想。另外几个问题是：庞加勒（H.Poincaré）猜想，关于三维球面的几何特征；理论计算机科学的P/NP问题，要求证明困难计算确实存在；代数几何中的霍吉（Hodge）猜想和伯奇/斯温纳顿-代尔（Birch/Swinnerton—Dyer）猜想；粘性流体动力学的纳维尔-斯托克斯（Navior—Stokes）方程的解是否存在；证明量子场论中的“质量间隙假设”。我想，我们到2050年会对那7个问题有

更多的认识，发现不同的结果。大概，庞加莱猜想那时还会悬着，P/NP问题将证明在形式上是不可确定的，霍吉猜想将被否定，伯奇/斯温纳顿-代尔猜想将被证明，纳维尔-斯托克斯方程在一定的奇异条件下没有解，质量间隙可能通过这样那样的方法解决，不过物理学家不会再对它感兴趣了。

700万美元的奖金不会把数学家引上新的路线。无论如何那是不可能的，因为数学家不像分子生物学家那么特别受金钱的诱惑。不过，它还是会达到一个目的，向圈外的人说明那7个问题有多重要——从而更一般地说明数学有多重要。我很愿意猜测这个信号会传到政府的基金管理部门，他们最终会认识到，把几十亿美元花在数学上，比花在新粒子加速器的零碎上或者别的什么庞大的生物学的“集邮”活动 [14] ，更能实质性地改变人类的生存状态，能产生更积极的影响。我想说，但我不会说。

P/NP问题是关于计算机的，却不是计算机所能解决的。它需要的是一个旧时的好思想。因为这个特殊的问题，计算机什么忙（即使是探索性的）也帮不了，不过它们能发挥别的作用，告诉数学家可能有什么猜想，然后数学家去寻求证明。我们现在越来越离不开这样的作用了。更重要的是，计算机将在许多证明里起着关键性的作用，这已经成为时下的趋势。在适当的程序下面，今天的计算机远不是60年代闹哄哄的数字机器所能比拟的，它们已经可以在严密的逻辑下充当我们证明问题的“助手”了。最有名的例子是阿佩尔（Kenneth Appel）和哈肯（Wolfgang Haken）1976年对四色定理的证明。那个定理最早是古斯里（Francis Guthrie）在1852年提出的，说的是平面上任何一幅地图只需要四种颜色就能区别出任意相邻的两个国家。证明的基本思想是把定理归结为一个程式化的识别过程，确认大约2000种特殊的地图（那是没有使用计算机发现的）具有一种特殊的数学性质。计算机进行了必要的计算，证明果然是那样的。

有些哲学家认为，计算机辅助证明在本质上完全不同于传统的证明，因为那样的计算不能由空手的人来检验。然而我们要问，为什么把人对证明的检验放在第一位呢？在这里，关键问题在于检验，而不在于什么样的检验主体。过去是人在检验，因为那时没有别的选择，但未来不一定还是人。首要的准则是检验的主体必须是可信的，谁如果不相信它，可以借助其他独立的主体来做出自己的检验。只要这样的条件满足

了，机器的判断与人的判断就是同样有效的。多数数学家希望计算机在投入运行的时候，能比人少犯计算或逻辑的错误。实际上，四色定理的历史中杂乱堆砌着人的错误。重要的还在于计算机程序的逻辑，在于机器是否真的朝着设计者的方向运行。这两件事情是可以独立检验的。“思想”部分现在还一直有人做——改变问题的形式，把它约化为大量的程式化的计算。然后，用计算机来算，抑或拿数学手册来帮助计算，在哲学上并没有什么不同。

在数学证明中让计算机充当探索的助手——就像生物学家的扫描隧道显微镜和基因排序机器一样——这种趋势将生根在2050年的数学中。那时会出现“虚拟的非现实”（VU）系统，数学家可以去“访问”抽象的概念结构，如非欧几何和大素数的范围，还可以任意操纵它们，几乎不费一点儿力气——跟我们今天拿计算器做算术一样。VU的原料已经备齐了，很快就能装配起来。软件工程师的需要将激发组合数学（有限的数学）的新发展。组合数学与几何学在今天难得发生相互作用，那时却能通过电路设计和逻辑函数的关系而结成亲密的伙伴。

在牛顿时代，数学问题的主要外部来源是天文学和力学，也就是自然科学。到2050年，更奇异的学科还会以同样的方式涌进数学。其中一个就是已经高度数学化了的量子物理学。今天，量子场论、几何学、拓扑学和代数学之间开始显现出新的令人惊奇的联系，跟着还会有更多。在未来50年里，量子场、超弦以及它们之外的各色理论所激发的新结构，将开出全新的代数和拓扑的天地。19世纪的数学家把传统的“实”数推广到“复”数，让“-1”有了平方根，也给数学带来了无限的生机。很快，数学的每一个领域都“复化”了：产生了与旧的实数的数学一样的硕果累累的复数的数学。量子化是21世纪的“复化”，我们将走进量子代数、量子拓扑、量子数论。

然而，影响更大、更剧烈的却是生命科学激发的数学：生物数学。尽管人类基因组计划成功了，它的结果却面临着新的现实问题。人们清楚地发现，DNA序列没有带来治病的良方，更没有为我们带来对生命的更深远的认识。在我们关于基因和生命的认识之间还存在着巨大的鸿沟。关于如何维护我们的生态系统，如珊瑚礁和热带雨林，基因序列没有告诉我们一点东西。人们曾经相信人类基因组有10万个基因，结果错了，只有34 000个。从基因走向蛋白质，那路线图比我们想象的复杂得多；实际上，也许根本没有那样的地图。基因是一个动态控制过程的一

部分，那个过程不仅制造蛋白质，还不断修正它们，使它们在演化的生命里，在生命历程的恰当时刻，找到自己恰当的位置。认识这个过程所需要的远不仅是一列DNA密码，而我们缺少的多数东西都是数学的。不过那将是一门新的数学，把生命生长动力学与DNA的分子信息过程融合起来的数学。DNA密码依然重要，但不是全部。新的生物数学可能是组合数学、分析学、几何学和信息学的奇异混合。当然，还要加上许多生物学。

在这个方向上，复杂系统的科学在不知不觉中蓬勃发展起来了。所谓复杂系统，指那些由大量以简单方式相互作用的、相对简单的组分（“中介”或者“实体”）所形成的系统。我们已经知道，表面上的简单是骗人的：从简单里能产生出高级的模式，即所谓的“突现现象”。例如，从人脑细胞的连通性产生出人的意识。到2050年，我们将拥有严格的关于突现现象和复杂系统的高级动力学的数学理论。它不仅会带来我们不曾梦想的概念，还将重新认识科学中的数学模型的局限。今天，复杂系统的研究主要在两个领域——生物和金融。例如，股票市场有许多中介，它们通过买卖股票相互影响。金融世界就从这样的相互影响中突现出来。金融和商务的数学将在革命中产生，它要抛弃现在流行的“线性”模型，带来数学结构更能准确反映现实世界的模型。

更引人瞩目的是，数学将走进整个人类活动的领域——走进社会、艺术甚至政治。然而，我们不会像今天的自然科学那样运用新的数学。在物理学中，数学用来表达定量的定律，而对现实世界的预言通常是大量计算的结果，在这些计算中，定律与产生的模式之间的联系不是我们的大脑所能跟踪的。例如，为了模拟飓风的巨大涡旋，我们需要写下几十亿个小区域暖湿气体的运动方程，然后通过大量的计算来解这些方程。另一个方法是从方程的一般结构（如对称性）导出涡旋的形态，不过现在还不够成熟。一种“涡旋的微积分”有可能把我们从无穷的数字纠缠中解放出来。更一般地说，我们有希望看到一个关于动力学模式形成的定性的、上下关联的理论出现在眼前。

最后，数学将帮助我们重新认识宇宙的模式——通过模式本身，而不是几十亿个跳动的数字，尽管模式是从那些数字中魔幻般产生出来的。

斯特瓦特（**Ian Stewart**）

斯特瓦特（Ian Stewart）因为对科学普及的杰出贡献而获得1995年皇家学会法拉第奖章。他为许多普及杂志写过大量关于数学的文章，如《发现》（Discovery ）、《新科学家》（New Scientist ）和《科学》（The Sciences ）。他为《科学美国人》（Scientifc American ）写过10年的“数学娱乐”专栏文章，也是《新科学家》杂志的数学顾问。另外，他与Jack Cohen合作写了《混沌的衰落》（Collapse of Chaos ）和《实在的幻境》（Figment of Reality ），还独自完成了《上帝掷骰子吗？》（Does God Play Dice ?）、《可怕的对称》（Fearful Symmetry ）、《从这里到无限》（From Here to Infnity ）、《大自然的数》（Nature’sNumber ）、《生命的另一个秘密》（Life’sOtherSecret ）和《更平直的世界》（Flatterland ）。

在文化的阴影里

B.古德温

Brain Goodwin

在历史的这个时刻展望未来是很困难的，同在1600年一样困难。那时，除了君王还在，西方的封建制度已经被彻底打碎了。在新兴的民族国家和新教派的双重打击下，神圣罗马帝国皇帝的权威灰飞烟灭了。在后来的几十年里，30年战争又把欧洲送进新的黑暗年代。[15] 那时，莎士比亚在用他的戏剧赞美复杂多变的人性，塑造从文艺复兴圣哲描绘的世界里走出的角色——在那样的世界里，天球的乐音表现着上天的和谐，爱的力量推动着世界的运行。伽利略在斜面上滚动他的圆柱，努力探索月亮和木星的奇特运动；不久，因为公开支持哥白尼，说地球确实在绕着太阳旋转（而不单是从数学观点看才那样），他受到了教会的惩罚。他犹豫着屈服了。[16]

培根高度赞扬了伽利略认识自然的方法的精神，但走进知识的科学方法还在黑暗当中。人类等级的教会戒条，还有莎士比亚魔幻的世界观，就是那时人们对万物秩序和人类在宇宙间的地位的基本认识。谁能

想到，在1600年阴影里的东西，50年后竟成为我们所谓现代的新文化方向的基础？伽利略的观察方法、测量方法和数学关系，为满目疮痍的文化带来了可靠的自然知识——那文化里一切可以信赖的东西，已经被其他知识体系的碎片和30年战争的蹂躏剥夺干净了。牛顿通过他的行星运动理论，以引力取代了那个推动世界的“爱的力量”，很快令所有怀疑者相信，科学才是通向认识的正路。

我们今天又来到一个文化的节点。这个时代是从近代走出来的，而我们已经看到，17世纪向近代的转变竟是那样意外，因此，我们今天说未来50年的事情，似乎是毫无意义的。不过，我们还是有办法为意外做好准备，即使不能预见将要发生什么，却可以让那转变更符合我们的心愿。这需要我们做一些跟本书其他文章相反的事情。我觉得应该把目光从未来50年移开，这样才好认识现在，并尽可能完整地感受它，特别感受那些处在阴影里的东西和正在开始走进光明的东西。这样，我们才可能摸索着走进一个新的未来，尽管我们不知道会走向何方。

看得见的

今天我们看到的，是科学、技术和商务的强大结盟，它开创了一个以预测、控制、创新、管理和扩张为首要原则基础的全球性的文化。这些原则的背后是理性和权力，培根曾倡导从这条道路去认识自然，用那些知识去砸碎人类的枷锁。这是一条康庄大道。通过科学知识的应用所创造的财富和工具，通过驱动资本主义发展的扩张机器，我们的确能够把人类从饥饿和贫穷中解放出来。然而，问题并没有像我们期待的那样得到解决：世界上还有许许多多（而且还在增加）的人生活在饥饿和穷困中；农业用地和自然资源正越来越多地遭到破坏；土地、海洋和空气的污染正影响着地球上所有的生命；矿物燃料的燃烧导致全球变暖，结果大气也变得更加混乱；物种正以自二叠纪和白垩纪末的生物大灭绝以来所不曾有过的速度走向灭绝；跨国组织的兴起，强化了无节制的全球货物和服务贸易，民族国家保护它们公民的能力越来越微弱。由于信息技术的惊人扩张，投资和资本运营的决策有时竟可能破坏市场的稳定，甚至瓦解政府。不断发展的气候混沌也广泛反映在政治混沌里，传统的说教似乎不能实现稳定与安全——那可是科学和技术（现代给我们的礼物）首先带给我们的东西。出人意料的是，我们今天又被拖进了黑暗年代，比30年战争更危险，因为今天面临的是全球的瓦解。

所有这些令人悲哀的事情都是看得见的。今天的这一幅伤心的景象，也许为我们带来及时的启示，也许提醒我们通过现代潮流的新的和谐与融合走向更加美好的未来。我不看好这两种可能，我想更需要认识我们目前状态下那些不容易看见却很可能突然涌现的东西。我的目的不是描绘未来的50年，而是要在未来图景还不明朗的今天，看看有什么朦胧出现的可能激发创造性活动的东西。

看不见的

我是一个科学家，为我们的未来做出巨大贡献的也可能是科学，所以，我要更多地反思隐藏在科学大厦外的可能发生作用的那些事情。我讲的第一个故事与伽利略在教会的经历有着共鸣，他当年被迫放弃了他支持的地球绕着太阳旋转的所谓异端邪说。

20世纪60年代，科学家和发明家洛夫洛克（James Lovelock）正在国家航空航天局（NASA）做着有关地外生命的事情。他发现，地球大气的组成使它在某种意义上不同于其他行星，这样，它可能向我们揭示出生命与非生命环境之间的某些深刻联系。他在科学杂志《自然》上撰文指出，生命并不是简单地适应它生根的行星所给定的环境，它还改变那些环境，让它们稳定下来从而使自己长期生存下去。这个观点得到了生物学证据的支持和光大，证据来自马各里斯（Lynn Margulis）对微生物改变行星条件的力量的研究；1974年，洛夫洛克和马各里斯联名在《大地》（*Tellus*）杂志发表文章，向科学世界公布了他们所谓的“盖亚假说”。这本是建立在牢固的事实基础上的科学，却披着古希腊大地女神的外衣。[17] 科学界从他们那听说了什么呢？他们把假说丢到外边的黑暗里去了。[18] 为什么？因为洛夫洛克和马各里斯违背了两个（而不是一个）正统的科学原则。第一，他们认为进化存在某些违背达尔文原理的基本东西。根据盖亚假说，生命并不简单服从地球给定的条件，而会改变那些条件来适应它们的生活。例如，微生物能改变大气组成（CO_2，CH_4，NH_3，O_2），这样温度就保持在生命能延续的一定范围内。于是，整个地球可以看作一个生命系统，它也像生命那样调整自己的活力。

第二，他们在假说里用了“盖亚”这个词，似乎说地球本身也是某种生命，而不是科学所认为的实现行星运动的盲目的机械过程的集合。大地女神的形象对积极的环境保护主义者来说是特别高大有力的，他们一

直在抗议对地球自然资源的掠夺，抗议对土地、海洋和空气的污染——如过度燃烧矿物燃料产生的污染。许多人在意识到我们对我们的地球做过什么并继续在那样做时，会感到伤心和愤怒，盖亚的形象成为那些情绪的一个焦点。

把盖亚请出来的洛夫洛克，因为假说里的那些邪说，结果被科学的教会驱逐出去了。关于第一点，他进行了有力的辩护。他承认达尔文的自然选择原理是一种进化机制，但他坚持认为生命除了适应地球环境的变化，本身也能改变地球的环境。今天科学界接受了这个原理，以“地球系统科学”这个名词来描绘洛夫洛克和马各里斯提出的那个更广阔的地球进化图景。不过，它的接受也付出了代价：洛夫洛克放弃了地球具有倾心眷顾生命的品性的观点——它听起来像万物有灵，是科学绝对禁止的东西；假如它真存在于我们文化的什么地方，那一定在最幽深的阴影里。

死的还是活的？

万物有灵论认为任何存在的东西从某种意义说都是有生命的，这不仅是科学驱逐的东西，也是我们文化的一般信仰体系之外的东西。最近，我跟圣塔菲歌剧团乐队的一个老练的长笛演奏者谈过话，告诉她我从一个纳瓦霍（Navajo）[19]长笛手那里知道，他为什么在演奏的时候带着七支长笛。他向我说明了每支笛子的不同音色和表现力。当我指着说像鸟叫的那支特别动听时，他又演奏了一遍，然后小声告诉我，在他们的文化里，指着笛子说话跟我们指着人说话一样，是很粗鲁的行为。说粗鲁是因为每支笛子都有名字和个性，一支笛子就是一个生命。歌剧团的乐手疑惑地望着我问：“他说的是真的？”尽管她很爱自己的笛子，可从来没有觉得它是活的。

为什么万物有灵的思想对西方的科学世界观有着那么大的威胁？难道有什么迹象说明科学的辩证法正开始让这样的观点重见阳光？假如真是这样，它的复苏在未来可能有什么重要意义呢？我们的科学坚信，作为宇宙基元的能量（或物质）是死的，没有任何感觉能力。伽利略从古希腊原子论者那里懂得了这样一个概念：除了原子和虚空，别无他物。以这样的眼光来认识事物，我们探索了世界的许多方面，对众多自然现象（不论死的还是活的）背后的各种过程的本质，树立了非凡的认识。这样的观念还为我们带来了数不清的令人难忘的技术。这条认识路线是

成功的，也是可靠的。根本的一点在于，大自然里能够量化、测量并完美地组织进描写各种行为法则的数学关系的那些东西，为我们提供了关于世界的唯一确定和客观的知识。当我们不用定量的语言，而用定性的语言来描述时，我们表达的是对所观察的事物的个人看法，如小水獭多么顽皮，风景多么美妙，朋友多么活泼。顽皮、美妙和活泼，都是性质而不能定量化的，所以它们不能作为可靠的现象描述的基础。它们也许有可以量化的方面，可以用来进行科学描述，但这些性质本身却是科学之外的东西，属于我们个人的主观经验。从这个观点看，说笛子是活的，有感觉的，指着它讲话就亵渎了它，这些说法都是没有意义的（比喻的意思除外）。笛子不仅是死的，而且什么经验也没有。

意识从哪里来？

最近走进科学议程的一个问题是关于意识的起源和本质。显然，意识的一个基本方面就是感觉，我们的感觉连同我们的思想，构成我们意识的内容。感觉可以是关于我们自己的，如我们感觉痛苦、快乐或者健康；也可以是关于外面世界的，如我们看到哭闹的小孩、受伤的动物或者垂死的老树。于是，在“意识从哪里来”的问题中，还有着另一个问题：感觉从哪里来？在科学中，我们的答案只能说，意识来自无意识物质在特殊的复杂和有序水平上的特殊的动力学组织，例如神经系统。我们的感觉是一种突现的性质，在产生它的事物中找不出一点儿可以称作感觉或者感觉能力的东西。我们面临的问题也在这里。

关于复杂系统的突现性质，我们有许多例子，它们都有着那种突现性质的某种形式的迹象。例如，蚂蚁成群结队来照顾它们的蚁后和卵，这种有节律的行为可以被认为是一种突现的性质。这是因为我们不能预言这种有序的行为会从一个个蚂蚁的行为（实际上是混沌的）和它们的相互刺激的相互作用中产生出来。不管怎么说，我们在蚂蚁群里观察到的就是那种有组织的活动，而且它还出现在模拟这种行为的计算机模型里。意外的有序总出现在以这种方式动态形成的系统中。

卵房里蚂蚁的集体节律的动力学前兆是什么呢？是一个个蚂蚁的活动（或不活动）的模式。从“混沌”一词的专门意义上说，蚂蚁的活动模式是混沌的：没有一个倾向的周期。然而，混沌是有节律的个体所构成的复杂模式，所以不难想象，当蚂蚁通过刺激相互影响时，就会突现一个倾向的节律。这里没有无中生有的奇迹。自然是和谐的，当我们看到

事情发生时，我们总能从系统的部分的行为和它们的相互作用模式来发现那现象的意义。同生物学一样，在固体物理学中发生的许多突现行为，也都是这样的。

然而，假如感觉是从毫无感觉迹象的物质中突现出来的，那我们实际上就是在无中生有。这对我来说就像一个奇迹。作为科学家，我更愿意在物质里添加一点儿某种形式的感觉或感觉能力，然后让它在以特殊方式组织的系统里放大——这个观点在哲学家的著作里有过广泛的探讨，如怀特海（Alfred North Whitehead，*Process and Reality*，1929），哈特肖恩（Charles Hartshorne，*Whitehead'sPhilosophy*，1972）和格里芬（Ray Griffin，*UnsnarlingtheWorld—Knot*：*Consciousness, Freedom, andtheMind—BodyProblem*，1998）。

质的科学

现在来看这种观点能带我们到哪儿。首先，物质里确实存在着感觉或者感觉能力，所以万物有灵论也不是那样不着边际。但是科学还有另外一面，正在开始改变的一面——它带着这种感觉和质的观点走得更远。与那种改变相关的是质的状态。现在有证据说明，当我们观察一个动物，说它紧张、狂暴或者孤独，实际上是在观察动物自己的感受，而不是把我们的感觉投射到动物的身上。证据为行为科学家维梅斯菲尔德（Francoise Wemelsfelder）和她伙伴的研究，说明不同的人在观察同样的动物时，有着高度一致的评价意见。科学就建立在这样的共识基础上——共识引出一个结论：我们看到的并不单是主观的，也是一种客观和真实的状态。所谓的"质的科学"就这样发展起来了，它是一种为了在评价中达成共识的方法，那些评价过去被科学家们认为是科学之外的事情。

正如我们看到的，今天的量的科学使我们能创造足够的满足地球居民需要的东西，但也在全球给生命留下迅速衰败的"质"。在现代科学的阴影里，同样能看到质的科学的成分，它让质的评价重回我们的日常生活，那里我们的判断不仅依赖于量，也同样依赖于质。质的回归连同我们对感觉的认识——认识到感觉不仅属于我们自己，也属于其他任何形式的自然物——可能极大地改变我们对科学技术的认识，改变我们团体和政治的作用。

这样剧烈的科学观念的转变，即使真的会发生，也不是一朝一夕的事情。它需要新式的基础教育，把自然科学与人文科学统一起来，让人们更完整；让社会的每一个成员来参与科学和技术的决策，让知识重新融入每一个负责的行为。于是，我们生活的整个时代将被后代认为是一个黑暗的年代，不过转变的种子已经在地球的阴影里萌芽了——也许可以说，盖亚正在那儿培育它们呢。

古德温（**Brain Goodwin**）

古德温（Brain Goodwin）是英国Schumacher学院（在Dartingdun, Devon）生物学教授，主持一个关于整体科学的宏大计划。他还是圣塔菲研究所的研究员。他是以下著作的作者：《细胞中的时间组织》（Temporal Organization in Cells ）、《细胞的解析生理学与生命形成》（Analytic Physiology of Cells and Developing Organisms ）、《美洲豹如何改变花纹：复杂性的演化》（How Leopard Changed Its Spots: The Evolution of Complexity ）、《形成与转变：生物学中的发生与相关原理》（Form and Transformation: Generative and Relational Principles in Biology ，与Gerry Webster合作）、《生命的信号：复杂性如何遍及生物学》（Signs of Life: How Complexity Pervade Biology ，与Richard Sole合作）。

替换大脑

M.豪瑟

Marc D. Hauser

想想下面古怪的事情：一只鸡长着鹌鹑的脑，像鹌鹑那样点头，像鸡一样叫；一个患了帕金森综合征坐在轮椅上的70岁老人，得到一点儿猪的大脑，立刻就能出去打高尔夫球了，看不出一点儿猪的影子。这不是亚当斯式[20]的科幻小说，而是科学事实。今天我们不但能在同类的个体间，还能在不同的物种间，交换脑组织。在未来50年，这样奇特的神经生物学将变革我们对脑的认识——脑如何在发展中形成，如何随时

间而进化。随着我们对脑的认识越来越多，我们最终将更好地理解做其他动物会是什么样子。不过，关于这场革命的科学和伦理学的后果，我们才刚开始考虑。

做另一种生命可能会是什么样子呢？这个问题是哲学家纳格尔（Thomas Nagel）在他那篇《做蝙蝠怎么样？》（*Philosophical Review* LXXXIII, October 1974，435—450）的著名文章里正式提出来的，主要说的是动物的“精神生活”，特别是关于它们的主观经历和感觉。在某些人看来，简直不可能还原那样的经历，至少凭我们今天的科学工具还做不到；另一些人则认为这尽管是一个难题，却在科学力所能及的范围内。如果有实例，这些问题说起来会更容易，下面我就举一个例子。

20世纪60年代中叶，有两个小组开始做恒河猴的实验，目的是想知道它们在看到别的猴子受到电击时会有什么反应。大约也在那时候，社会心理学家米尔格拉姆（Stanley Milgram）正开始考察人的行为，看他们对权威的反应——特别是，当某个权威人物让他们去电击另一个人的时候，他们是否服从他。在其中的一个猴子实验里，一只猴子被训练去拉操作杆，然后才得到一天的食物。等它学会以后，把另一只猴子关进隔壁的笼子里。这时候，第一只猴子拉动操作杆，就会对新来的猴子产生很强的电击。奇怪的是，它不但停止了拉操作杆的动作，而且好几天都不再那么做，尽管因此而失去了吃的东西。它虽然饿着肚子，邻居却不会遭电击了。假如邻居不是一只陌生的猴子或其他什么动物（如兔子），而是它熟悉的“笼友”，那猴子似乎更不愿意去拉动操作杆。最后，经历过那种处境或挨过电击的猴子，会比没有那种经历的猴子更远离操作杆。

从截然相反的人类实验的结果看，恒河猴实验的结果是特别令人惊讶的。在1983年的《服从权威》（*Obedience to Authority* ）一书里，米尔格拉姆生动描述了他的人类实验。当一个权威人物（如一个穿着白大褂的实验家）命令实验者拉动操作杆电击别人的时候，实验者会不停地那么做，而不管那人对“电击”产生多么强烈的反应——当然，那是一个演员，并没有真的危险。假如哪个火星人降临地球，看到这样的两个实验，他只能得出结论说，恒河猴有同情心，而人类没有。恒河猴似乎懂得处在别人痛苦的境地会是什么样子，而人类要么是不懂，要么是漠不关心。当然，我们知道人是能够关心同情别人的，也能设身处地想着那些经历情感波澜的人。爱略特（George Eliot）的动人小说《亚当·比

德》（*Adam Bede* ）的读者，能很好地体会亚当见海蒂时的感觉，那是一种不求回报的爱。下面一段文字大概说明了这一点：

> 那羞怯令亚当心跳，生出从未有过的幸福。她以前看见他时从不脸红。“我吓着你了。”他说，隐约觉得说什么都无所谓了，因为海蒂似乎跟他有同样的感觉…… [21]

恒河猴关心和同情其他伙伴的状态，这个结果很诱人，而且实验似乎也证实了这一点。不过还有另外的解释：也许拉操作杆的猴子发现被电击的伙伴表现出厌恶的情绪，在被人厌恶的情形下，人们总会停止正在进行的事情；也许猴子担心自己以后会处于那样的境地，受不懂怜惜的家伙摆布。假如是这样，猴子不拉操作杆就不是因为同情，而是因为自己的利益。不论哪种解释是正确的，这些实验都揭示了猴子对社会环境的敏感反应：说明了猴子有情感和目标，而且可以照着它们行动。利用这些信息，我们可以把研究动物的思维与提高动物的福利联系起来。

怎样才能发现动物的需要、愿望和目标呢？如果我们想切实地关心它们，这是很重要的信息。假如洛夫汀（Hugh Lofting）的杜里特医生不是小说里的人物，我们能跟动物对话，那该有多好！ [22] 不过，我们确实有很好的办法来代替那样的对话。从仔细观察动物的典型行为开始，然后借助经济学的手段来看它愿意为它想要的东西付出多少代价。我们来看看最近对农家饲养水貂的研究。

养水貂的农家相信他们的动物过着舒适的生活。在这里，“舒适”的意思多少有点儿像我们所谓的“丰衣足食的小康生活”。不同意水貂的这种生活观的人，不相信“小康”就意味着有足够的吃喝，没有大的病痛。空洞的理论解决不了这场争论，还要拿干脆的实验来说话。剑桥大学生物学家玛松（Georgia Mason）和她的同事把水貂放置在一个个笼子里，模仿农家的饲养条件：水、食物和睡觉的窝。他们认为，所有的动物都喜欢寻求享乐，向往好的东西，躲避不好的东西。在这个认识基础上，他们还为每一只水貂准备了七个可以选择的安乐窝，每一个分别配备着某种独特的设施：装满水的小池塘、抬高的平台、新奇古怪的东西、两个小窝、一条小通道、几样玩具、额外的活动空间。为了走进那些笼子，水貂必须推开相应的门。在接连的几天里，每一道门都系着重物，不容易打开。实验还假借了保守的经济学，每个人需要为他想要的东西付出代价。这些实验的重要基础是，我们直觉地认为，动物不但可能为

想要的东西付出更多，关键的是，它们还可能为需要的东西付出更多。

把水貂从原来的笼子里放出来，它们不约而同都选择了有小池塘的新家，多数时间都待在里头，并为它付出了最大的代价。而且，假如剥夺了它们的水池，通过测量紧张荷尔蒙（即大家熟知的皮质醇）水平可以发现，它们像失去了食物一样难过。水貂们想要什么呢？水池。为什么？因为在自然的栖息环境下，它们就喜欢长时间待在水里游泳，在水里寻找吃的。结果，为了让水貂过上“小康生活”，饲养它们的农家就得把它们可怜的“生活费”拿出来为它们买小水池。没有水池的水貂跟被剥夺了食物的水貂一样难过。没有哪个仁厚的农家想过去剥夺水貂的食物，又怎么会去剥夺它们的水池呢？那样做既不经济，也不道德。

恒河猴和水貂实验说明科学家如何能发现动物的感觉和需要，又如何更好地发挥这些知识的实践作用。不过，从当前的基因和脑科学的发展看，我们描述的技术也太粗糙了。既然我们不但能移植或替换部分大脑，还能通过增加或删除基因来改变动物的基因组，那么我们能提出和回答的问题也就非常广泛了，而潜在的道德伦理的难题也同样广泛。想想最近出现的聪明的“道奇”鼠[23]。这些小老鼠是遗传工程的产儿，它们被注入了更多的NR2B基因，那个基因对记忆形成起着重要作用。人们相信拥有更多这种基因的老鼠要比对照实验的普通老鼠聪明，因为它们能更快地学会选择对象，感应厌恶的东西，发现隐蔽的陷阱。这些增强的本领是否构成智能的基元，当然还可以争论；不管怎么说，结果总还是揭示了一些似乎通过基因手段的作用而表现的不同。如果对更高级的认知功能的基因感兴趣，这些结果是很令人欣喜的。它们不但体现了技术进步的力量，还展现了那些可能有实际应用（特别是对人类疾病的治疗方法）的基因工程。例如，在大脑中增加与记忆相关的感受器的数量，可能从理论上逆转像阿尔茨海默症[24]患者那样的毁灭性的记忆丧失。

但是，道奇鼠的发现却因为另一个实验的结果而多少有些扫兴——那个实验揭示了操纵基因和大脑的潜在危险。在道奇鼠出生两年后，科学界面前意外出现了一个由聪明带来的副产品。那个副产品正好应了一句运动员的格言：“不经风雨，哪来彩虹？”跟普通伙伴不同的是，道奇鼠对巨大的痛苦有着更长时间的体会。这个结果有着重要的意义。正如遗传学家勒文廷（Richard Lewontin）在《三螺旋：基因、生命和环境》（本书也是他对人类基因组计划的批评）所指出的，我们要避免对

基因与行为的因果关系做出天真的结论，否则我们不可能分辨复杂的基因组的背景和基因所处的环境背景。那里是基因的丛林，当一个基因被清除、被取代或者被复制，我们对结果只能做出经验性的（也就是统计的）猜测。这并不是说操纵基因和大脑没有价值，相反，那样的技术很可能开出一片新发现和新认识的天地。不过，伴随着这些发现，我们必须准备面临意想不到的复杂和困难。

我常让学生做下面的思想实验：假如你有机会经历可逆的大脑移植（就是说，你还可以完整地要回你原来的那部分大脑），接受某个动物大脑的某个特别的部分，你愿意选择什么动物、选择哪个部分呢？近年来，我的学生们列出的前三个选择是：狗的嗅球，蝙蝠的听觉皮层，鹰的视觉神经。这个思想实验藏着一个不易觉察的陷阱。尽管在技术上可以移植那些大脑皮层区域，但是要真做到像狗那样闻，像蝙蝠那样听，像老鹰那样看，还需要一点儿别的东西。别的东西指的是一个解析系统（附带着些外在的器官，如狗的鼻子，蝙蝠的雷达天线似的耳朵，鹰的凹陷的眼睛）。人装上新的狗嗅觉系统，可以像狗那样在百米之外闻出消防栓里几毫摩尔的尿，但还是以人的方式解释那个气味。也许我们会因为气味太强烈而觉得可怕——以前还从来没人受过那么强烈的刺激。

我想强调一下我们大脑行为中解释方面的重要性，因为它常常被忽略了。我们大概可以通过一个哲学悖论和一部恐怖电影来说明这一点。逻辑上关于恒等的理论说，任意两个具有多个部分的物体x和y，如果x的每个部分都属于y，y的每个部分都属于x，那么x=y。恒等概念遭遇的经典挑战是忒修斯和他的雅典水手的船。船出发的时候是新的，经过长时间的风吹雨打，船板坏了，水手换了新的。到航行结束的时候，原来的甲板和设施都换过了。问题是，航行到终点的船还是启航的那一只吗？还是忒修斯的那只船吗？[25] 回答问题之前，我们来看波兰斯基的电影《房客》。[26] 波兰斯基扮演的是一个性情温和的档案管理员，住在巴黎的一家公寓里。以前的房客曾经自杀过，这令他常处在疯狂的幻觉中，于是有了下面的两句拗口的关于自我本性的独白："如果我斩断我的手臂，我说'我和我的手臂'，但是，如果我砍掉我的头，我还说'我和我的头'或者'我和我的身体'吗？"上面两个例子特别表现了与解释相关的困难。假如我们割除某人的嗅觉系统而拿另一个人的或狗的嗅觉来代替，我们并不会改变那人的本性，只是改变了他感觉气味的方式（特别是跟狗交换嗅觉时）。换了新嗅觉的人对气味的解释还是原来那样。但是，当大脑被替换了，我们就得逐个案例地提出同一性的问题。神经

学家蒂梅修（Antonio Damasio）在他最近关于意识研究的《感觉发生的事情》（*The Feeling of What Happens* ）一书中明确指出，大脑的不同部位，对感觉发生在自己身上的事情，有着不同的影响。某些可以替换的部位很可能产生巨大的个性改变，著名的盖奇（Phineas Gage）的例子很清楚地说明了这一点。盖奇原是个勤劳的受人尊敬的人，后来脑前额皮层被损伤，失去了一切精神判断能力，变得大家都不认识了。[27]

为了把“思想替换”的问题推得更远，我们可以在来自神经科学世界的一些令人惊奇的新结果的基础上，做另一个思想实验。神经生物学家尼可勒里斯（Miguel Nicolelis）和他的同事们设法记录了鹰猴大脑的几百个神经元的电流，用这个信号来驱动机器人的手臂。这听起来简直是玩游戏，但不是那样的。它说明我们可以在一定水平上弄清神经密码的意义，认识它如何调节行为。现在让我们来想象，我们可以从任何动物的大脑“下载”神经信号，建立一个硬驱的思想图书馆，记录动物与世界发生作用时的思想状况。在动物吃饭、睡觉、梳洗、交配和相互沟通的时候，我们可以读出它们在想什么。在一定程度上，我们也许能更深切地体会，当我们成为它们的时候，会是什么样子。我们也许是爱偷窥的“智人”（Homo Sapiens）；我们甚至可以拿自己的脑电波来跟它们进行对比，从而体验前所未有的人与动物的和谐——那显然是虚拟现实游戏的终点。

这是一些奇妙的思想实验。在未来50年里，我们需要的技术都能实现，尽管不会有人以那样奇怪的方式去运用它们。我们将知道多少关于大脑的东西——不论我们的还是那些能思想的动物的——想起来就令人激动不已；而我们的技术正把我们引向一个陌生的充满了模糊的道德问题的世界，这又令人担忧。假如我们替换一部分大脑，开启或关闭某些基因，那后果该由谁来负责？科学家，医生，还是那献出了大脑让人能更好生活的动物？假如干细胞研究被批准了，大脑的不同部位可以独立培养，是不是任何人都能做替换？科学依靠献身科学的人的创造性能量，就这一点而言，思想潮流会倾向激进的甚至危险的探索，但科学家也必须认识到他们的行为（也包括非人类动物的研究）的潜在的伦理学后果。萧伯纳（George Bernard Shaw）在《巴巴拉少校》里思索，正确与错误的秘密“曾令所有哲学家伤神，令所有律师疑惑，令所有生意人糊涂，而令多数艺术家崩溃”。也许他还应该加上科学家，科学家必然还要继续努力去分别，哪些是“正确的”结果，哪些是“应该的”结论。

豪瑟（**Marc D.Hauser**）

豪瑟（Marc D.Hauser）是认知神经科学家，哈佛大学心理学系神经科学项目和“思维-大脑-行为行动”教授。他写过《交流的进化》（The Evolution of Communication ）、《动物交流设计》（The Design of Animal Communication ，与M.konoshi合作）和《野性思维：动物真的在想什么》（Wild Minds：What Animals Really Think ）。

儿童会告诉科学家什么

A.戈普尼克

Alison Gopnik

1997年，在国家航空航天局（NASA）工作的空间科学家们发现了怎样通过分析从火星岩石反射的光线来确定火星是否存在过水。水会在岩石上留下碳的痕迹，这将影响岩石反射的光谱。科学家可以反过来从光的数据寻找碳，然后根据碳来确定水的存在。那年，在几千米以外的伯克利幼儿园，一个名叫克温（Kevin）的4岁小男孩儿也同样兴奋地发现了一台新机器是如何运转的。在机器上放一定组合的积木块儿（不能是别的东西），它就能奏响音乐。小克温根据这个事实反过来推测哪些木块儿能让机器响起来，然后他用这个发现来让它奏出音乐。在未来50年，我们将逐渐懂得克温和NASA的火箭专家们如何能够做出这些惊人的发现。问题的答案还将使我们换一种方式去思索科学、儿童、大脑，也许还有基因。

人类对周围的世界有许多了解。我们知道岩石、波浪、烤炉；知道兔子、棕榈树、牵牛花；知道父母、孩子、正牙医生——还有数不清说不完的人物、动物和植物。我们的知识大体说来是很准确的：我们能很好地预言烤炉、牵牛花和牙医的活动。当我们每天摁下“烘烤”按钮，在花园添加“神奇生长灵” [28] ，或者预约病人的时候，都在运用这些预言。出生的时候我们什么都不懂，但不管怎么说，我们还是学会了。

我们也学日常生活以外的事物，如火星岩石、病毒和神经元。这些知识同样是非常准确的——足以让我们控制或至少减缓如天花和抑郁症等古老的疾病，更不用说秃顶、阳痿和偏头痛。

但我们怎么能懂那么多呢？毕竟，我们直接从世界得到的信息不过是打在我们视网膜上的一个个无穷小的光子和在我们耳膜振动的无规则的气流。从那样一点有限的而且显然毫不相干的信息，怎么可能得到真理呢？“真理”说起来像一个堂皇的形而上学的概念，其实我们都知道很多日常生活中的真理：热烤熟了面包，水滋润了花朵，病人失约惹恼了正牙医生。从心理学的观点看，这样的一些知识与理论物理学或天文学的知识是同样令人惊奇和疑惑的。一种类型的物理客体，如顶着颗头颅的一个皮囊，与另外类型的物理客体，如烤炉、牵牛花和正牙医生，两者之间的系列相互作用，如何能让一个去学会认识另一个呢？

近50年的发展心理学使这个问题变得更加疑惑。因为有了新技术，我们能比从前更多了解儿童的思维。看来，婴幼儿知道的和学习的都超出了我们从前的想象。到了三四岁的时候，他们已经大概知道了世界是怎么回事。一个小孩子，不会读，也不会写，连话都说不清楚，怎么能那么快地学会那么多的东西呢？这是学习的理论需要解释的问题。我们的学习能力不能仅仅归功于教育、训练或者什么专门的社会机构，它似乎还是我们人类本性的一个基本组成部分。

在过去的50年里，认知科学告诉了我们很多东西：我们关于世界的知识是什么样的；我们如何应用那些知识；那些知识是如何印入我们大脑的；发展认知科学还告诉我们知识如何随我们年龄的增长而改变。但我们还是不知道那些知识从哪里来，又如何能给我们带来我们外在世界的真实图景。学习的问题，与意识和浪漫的爱的问题一样，已经写进了认知科学教科书里“未解之谜”的一章。关于意识问题，我不相信我们能在50年里得到更多的认识，浪漫的爱就更难说了。不过，我想我们能实在地接近一个关于学习的科学解释。

我们能在另一个迥然不同的认知科学的领域找到那种解释的模型：人类的视觉。视觉问题是这样的：获取进入眼睛的光的模式，然后把那种信息转化为在空间运动的物体的精确表示。我们是如何解决这个问题的呢？进入眼睛的光如何与空间事物发生联系，人们似乎隐约做过一般的假设。例如，我们似乎不自觉地假定进入视网膜的光是三维世界的二

维投影，我们就用这样的假设来解决视觉问题。我们从不认为我们生活在平直的世界，尽管那在逻辑上是可能的。实际上，婴儿似乎生来就相信那一点，例如，小宝宝在看见东西向他们逼近的时候，总会向后退缩。

不过真正有趣的还不在于我们认识了视觉的这件事，而在于通过假定那样的事实，我们就能发现许多意想不到的新的事实。我们不自觉地假定，视网膜的图像是三维物体的二维投影。根据这个一般的认识，我可以猜测现在我视网膜上的那个特殊图像，一定来自一根细棍连结的两个圆盘，它正以一个特别奇怪的角度躺在地板上。知道了这个事实，我可以解决我在现实中遇到的麻烦：没完没了地到处寻找我读书用的眼镜。

当然，这样的假定有时也会把我们引入歧途——特别是当某个邪恶的心理学家在制造幻觉的时候。不过，这些假定往往都是正确的，能让我们正确认识外在的世界是什么样子。

但是大脑怎么能做出假定呢？当大脑（或其他计算机器）接收一定输入时，我谈的那些假定会转化为对输出结果的约束。当我的视网膜以特别方式“亮”起来时，只有某些神经元（而不是所有的）能接着亮下去。神经科学家在动物注视一样东西的时候，可以记录它的视觉皮层上某些特殊细胞的输出结果，然后构造出一个网络图像。神经学研究说明了那些约束是如何发生作用的，那些计算又是如何在大脑中实现的。

在视觉科学中，不同的学科令人惊讶地走到一起来了。心理学家告诉我们从什么样的视觉信息形成什么样的物体表象；告诉我们什么模式的光打在视网膜上会产生什么样的感觉——问题就这样确定下来。然后，数学家告诉我们如何才能通过确定关于物体与光联系的非常一般的假定来解决那个问题。计算机科学家告诉我们那些解如何作为实际的物理机器运行的约束而实现。而神经科学家告诉我们那些解又如何在我们头盖骨下的那个特殊机器里实现。

同样的路线也有助于认识我们是怎样学习的：那就是，确定儿童和成人解决的问题，在一定假设下用数学得出那些问题的可能解，看那些解如何在计算机里实现，然后看它们最终如何在我们的大脑中实现。最近，来自不同学科的——科学哲学、人工智能、统计学和发展心理学

——关于学习的新思想，也同样地走到一起来了。在未来50年里，这个认识的融合将产生一个羽翼丰满的关于我们学习的科学理论。

我们就从这个问题说起——它至少是出发点的问题之一。我们如何认识世界的因果结构——事物如何运动，一个事件如何引发其他的事件？在任何科学实践中，这当然是一个重要问题，而对幼小的儿童来说，它也是重要的。发展心理学已经证明，儿童懂得许多有关因果关系的事情。到三四岁的时候，他们就跟大人一样知道了烤炉、牵牛花和人。五岁的孩子比三岁的孩子知道得多，而七岁的孩子知道得更多。儿童跟科学家一样，似乎很容易学会新的因果事实。

不过，因果知识也代表了我们的经历与我们的学问之间的一道深广的鸿沟。哲学家休谟（David Hume）最早阐发了这个问题。我们所看到的不过是事件之间的可能事件。一类事件可能总是伴随着另一类事件发生，但我们怎么知道一个引发另一个呢？在现实生活中，因果关系很少只涉及两个事件的，可能有几十个事件以复杂的方式因果地联系在一起。在现实生活中，一个事件实际上通常不会总跟着另一个事件，而且我们也并不总是知道两个事件中的哪一个先发生。这样的不确定性和复杂性使我们日常的因果问题显得更加复杂。是烤炉电阻丝的烟烤熟了面包片，还是灼热的面包屑使烤炉的电阻丝冒烟？或者，我们把温度调得太高了，它在烤熟面包片的同时也使电阻丝冒烟？我们能看到的只是同时出现的一片混乱。

有什么办法来清理那一堆混乱吗？直观地说，我们可以做两件事情。我们可以做一系列实验：例如，我们把温度旋钮定在一个很高的温度，但烤炉里没有面包；或者，我们把灼热的面包屑洒在电阻丝上，但让温度保持在很低的状态。假如做不了实验，我们可以仔细去观察，确定电阻丝在什么时候冒烟，什么时候不冒烟。也许，不管有没有灼热的面包屑，它只在高温下才冒烟。或者，不管温度多高，它只在有灼热的面包屑时才冒烟。

当我们做这些实验或者观察的时候，我们也在假定系列之间的可能事件的发生模式是如何与它们之间的因果关系相联系的——正如我们假定二维视网膜图像如何与三维物体相联系。我们从没想过我们生活在平直的世界，同样，我们也没想过生活在没有原因的世界。当然，同视觉的情形一样，休谟的魔鬼也可能以某种方式安排那些可能事件来欺骗我

们。但是，我们的进步却是因为我们不相信有那样的魔鬼——用爱因斯坦的话说，上帝虽然狡猾，却不邪恶。[29]

卡内基—梅隆（Carnegie—Mellon）大学格里默（Clark Glymour）领导的一群科学哲学家与加州大学洛杉矶分校（UCLA）的计算机科学家皮尔（Judea Pearl）和他的同事们，开始创立一种数学形式来帮助我们跳出直觉，以严格的方式表述那些假定。我们可以借助所谓的定向无圈图（通常也叫贝叶斯网络图）[30] 来思考因果关系。这些图告诉我们一个变量（如面包的状态）如何影响另一个变量（如热电阻丝的状态）。无圈图背后的基本假定是，如果一个事件引发另一个事件，那么当那个变量的数值改变时，另一个变量的数值也可能发生改变。如果面包屑导致电阻丝冒烟，那么烤炉里面包屑的存在应该更可能引起冒烟。我们可以用连接变量的箭头来表示这些因果关系。贝叶斯网络图假定了一些简单而普遍的因果关系模式（也就是箭头的模式）是如何与变量间的可能事件的模式相联系的。在温度旋钮、灼热的面包和冒烟的电阻丝之间，我们能画出三种不同的图像，相应于我们说过的三种因果假设：

A.温度旋钮＞灼热的面包＞冒烟的电阻丝

温度旋钮使面包发热，灼热的面包使电阻丝冒烟。

B.温度旋钮＞冒烟的电阻丝＞灼热的面包

温度旋钮使电阻丝冒烟，冒烟的电阻丝使面包发热。

C.冒烟的电阻丝＜温度旋钮＞灼热的面包

温度旋钮分别使电阻丝冒烟，使面包发热。

在一定的可能事件和因果性的基本假定下，每个因果结构对变量间可能事件的模式有着不同的意义。我们能通过实验和观察做出正确结论，也是因为这一点。举例来说，假如A是对的，我们把面包拿走，就不会看到温度与电阻丝之间有什么关系了；假如B或C是对的，我们还能看到那种关系；假如我们把旋钮定在低温，另外独立地加热面包，那么在A的情形，电阻丝会冒烟，而在B或C的情形则不会；假如把旋钮调到高温，但不让电阻丝冒烟，那么面包在A或C的情形下会发热，而在B则不会。同样，即使我们自己不做实验，不同的因果结构也能使我们在

变量间看到可能事件发生的不同模式。数学的研究使我们具体认识了可能事件和原因之间的所有联系，它们的结构比我上面讲的那些要复杂得多。

这样的数学为我们提供了一种因果的逻辑。古典的演绎逻辑从几个关于推理的基本假定出发，把那些假定转化为一种数学方法，从而根据真的前提导出正确的结论。新的因果逻辑则做出几个关于因果性的基本假定，然后提出一个系统的方法来根据观察和实验导出正确的因果关系的结论。

计算机科学家已经开始把这种抽象的数学转化为能向现实世界学习的计算机程序。计算机与人之间的一大区别在于，计算机程序只能做我们首先要它做的。真正的图灵试验——检验计算机是否像人——不仅要求计算机像成人那样做事，还要求它学会根据儿童的经验来做这些事情。[31]

在一定的可能事件的数据模式下面，计算机科学家将那些数学假定转换为对计算机可能产生的因果图模式的约束。根据这样的数学新思想，在NASA工作的计算机科学家可以设计一些程序让机器人单纯根据光谱仪的数据来认识火星岩石的组成，而不需要向地球上的专家请教。

上面所说的一切，离我们出发点的问题似乎太远：我们想知道寻常百姓——特别是普通的儿童——到底是怎么学习的，而不是只想知道高明的科学家、统计学家和计算机专家是怎么学习的。不过开始出现了一些证据，说明所有的学习者都可能对因果性与可能事件之间的关系做出相同的数学假定。心理学家在研究普通成人解决因果问题的方式时，已经独立发现了与哲学研究者相同的数学模型。

心理学家发现，即使只有两岁的儿童，也在运用同样的因果逻辑。我们可以向孩子们拿出跟烤炉相似的东西——一台以复杂和神秘方式发生某些事情的机器。有时我们把机器的一些特殊行为模式告诉小孩，有时我们让他们自己做实验去发现那些证据。然后我们看他们是否能发现机器是如何运转的。令人惊奇的是，孩子们能完全以模型所预言的方式，根据事实做出正确的因果结论——小孩儿真是天生的火箭科学家。当然，小孩跟科学家不同，他们一点儿也不知道自己是怎么得出那些结论的。

在未来的50年，一旦知道小孩和大人正在做什么计算，我们就能深入他们的大脑，看他们如何计算。随着脑图像技术越来越精确，对计算的认识越来越多，我们将开始去认识我们的大脑是如何设计来完成那些计算的。那个问题的答案也许关系着即将来临的神经科学的伟大突破。大脑最重要的事情是它为顺应环境条件而发生改变的能力，不过大脑的这些方面，正是我们知道得最少的。这就像我们对于死亡心脏的解剖结构什么都知道了，但是对于一颗活的沸腾着热血的心却几乎一无所知一样。大脑是最重要的学习器官，如果知道了学习是怎么进行的，我们就会知道大脑是怎么工作的。

还有，在接下来的50年中，关于思维和大脑如何学习的问题，答案可能会牵涉到更一般的发展的问题，即关于形态发展的问题。新千年的另一个未解的难题是，DNA的指令怎么能使一个简单的受精卵变成一个高度复杂的小生命？过去几年的遗传学研究告诉我们，基因组不可能是产生生命的一套具体的指令集合：它不是生命的蓝图。那么它又是如何发挥作用的呢？基因的作用并不太在于直接决定细胞做什么，而是更多地在细胞环境里激发一连串的因果行为，最终以可以预料的方式来影响细胞。例如，决定性形态的基因产生睾丸激素，然后通过睾丸激素以一定的方式作用于生命。有时候“魔鬼”可能出现，环境最终会变得跟基因“希望的”不一样，系统可能发生错误。但环境通常是可以预料的，基因组根据那些可以预料的事情来产生复杂的生命。

为了更好理解，我们可以把DNA指令看成一套密码，它隐含着细胞（不论在生命诞生以前还是以后）与环境（主要是其他细胞）相互作用的一般性假定。在心理学的情形，通过假定我们与环境的关系，我们建立了能很好适应那种环境的复杂结构。至少，我们相信这种一般性的方法也适用于生物学的情形。

然而，一个统一的学习理论的最大成功，也许在于说明最杰出的科学家与最普通的儿童原来参与了相同的事业。在20世纪的尽头，知识就像封建经济的土地和工业经济的资本，成了最有价值的货币。学习的新理论应该告诉我们，知识的获得，不仅仅是为了在今天残酷的竞争中赢得什么显贵的地位。从本来的（而不仅是修辞的）意义上说，知识是全人类与生俱来的权利。

戈普尼克（**Alison Gopnik**）

戈普尼克（Alison Gopnik）是伯克利加利福尼亚大学心理学教授。她是国际儿童学习领域的领头人，也是利用发展心理学帮助解决古老哲学问题的认知心理学家之一。她同Andrew Meltzoff合作写了《语言、思想和理论》（Words, Thoughts, and Theories），与Patricia Kuhl和Andrew Meltzoff合作写了《婴儿床上的科学家：思维、大脑和儿童是怎么学习的》（The Scientist in the Crib: Minds, Brains, and How Children Learn ）。

走向道德发展的理论

P.布隆

Paul Bloom

大学生上第一堂心理学课（通常是一般的概论）的时候，往往惊讶它是那么无聊。他们满怀希望走进课堂是为了学习他们是如何进行思维的；在他们模糊的心目中，心理学是关于梦、意识、邪恶、疯狂和爱的学问。学期结束离开课堂的时候，他们只隐约记得什么抑制性突触，巴甫洛夫（Pavlov）的狗，斯金纳（Skinner）的老鼠 [32] ，还有可笑而恼人的社会心理学实验，最新的精神疾病分类法。但是，他们没有得到曾令他们兴奋的问题的答案——更糟糕的是，他们觉得连问那样问题的人都没有。

10年前大概是那样的，不过现在不同了，有许多迹象表明，到21世纪中叶时，心理学可能包罗万象，一点儿也不会无聊。它将拥有广阔的范围和丰富的理论。它将把来自不同领域的发现、方法和思想，包括进化生物学、文化人类学和思维哲学，融合在一起。换句话说，今后50年的心理学会更像它100多年前在19世纪末的那个样子。

那可是激动人心的年代。在1859年出版的《物种起源》的最后，达尔文（Charles Darwin）写道："我看到了将来更加重要得多的开放的研究领域。心理学将稳固地建立在一个新的基础之上：每一智力和智能都是通过逐级的演进而必然获得的。"（20年前，他就在自己的笔记本里写道，"任何理解狒狒的人都能比洛克 [33] 更接近形而上学。"）达尔文

试着在后来的两本书里实现他的这个愿望：一本是《人类的由来及其性选择》，主要是为了解释人类与其他动物之间的心理学差别；一本是一年以后的《人和动物的情绪表达》，对情绪表达的心理学和生理学的研究者来说，今天仍然是一部十分重要的著作。

1890年，詹姆斯（William James）发表了《心理学原理》，总结了那个时代最好的科学，提出了野心勃勃而充满个性的观点。[34] 那时候，弗洛伊德（Sigmund Freud）正在另一个很不相同的方向上开展研究。尽管弗洛伊德对人性有着持久的影响，在当代心理学（不论临床的还是实验的）中的地位却不那么稳定，也许他还能出现在当代心理学的导论性课程里，那不过是一个历史趣味的话题（假如不是嘲讽的对象的话）。但他那广博的见识、激情和野心却是异乎寻常的。我们只看一个例子：1899年出版的《梦的解析》的头几句话是这样说的："在后面的篇章，我将证明存在一种允许我们解释梦的心理学技术，运用那种技术以后，每一个梦都将显现一种有意义的精神结构，可以在我们醒着的生活中分辨出它们的位置。"弗洛伊德的宏大研究纲领的精神——在无意识过程的相互作用的基础上统一精神科学的目标——与达尔文和詹姆斯的工作是遥相呼应的。

后来的百年里，心理学从其他领域（特别是哲学和进化生物学）分离出来，成为独立的学问。心理学有意识地、自觉地寻求成为一门科学。努力的结果之一是20世纪席卷美国心理学的行为学派的运动。行为学派排斥詹姆斯关于心理学是"精神生活的科学"的观点，提出只有可以观察的行为才能进行客观的研究，几乎所有的行为都是学习的结果，不论距离多远的物种（如人与老鼠），在学习上都不存在原则性的差别。

今天，行为主义像任何陈腐的理论一样死了，主要是因为所有的那些前提都被证明是错误的。尽管从它的研究纲领中涌现过一些重要方法——为了测试婴儿和老鼠那样不能清晰发音的动物的能力，这些方法是很有用的——却几乎没有产生过与人的研究相关的影响久远的发现。席卷今天心理学的运动是为精神生活提供计算分析的认知心理学——最近的分析方法是平行分散处理的动力学，也就是神经网络。这个研究纲领获得了空前的成功，但成功主要局限于那些容易在计算机上模拟的领域。于是，我们有了许多下棋、推理、事物识别、语言分析和不同记忆形式的研究。而情绪、性行为、动机、个性等问题却被推向应用性更强的领域，如临床心理学。

因为与其他领域越来越多的交流，这一切正在发生改变。心理学中某些最具影响的思想是从外面的领域走进来的，这绝不是偶然的——它们来自哲学家，如丹内特（Daniel Dennett）和弗多（Jerry Fodor）；来自进化论理论家，如哈密尔顿（William Hamilton）和特里弗（Robert Trivers）；也来自经济学家、人类学家和语言学家。对心理学领域最具影响的学者之一无疑是语言学家乔姆斯基[35]，他1959年对斯金纳《言语行为》（*Verbal Behavior*）的批判，是对行为学派运动的致命打击。

特别有趣的跨学科交流是与进化生物学的联系。最近几年里，越来越多的人接受了达尔文的思想：大脑跟其他生物器官一样也经历过自然选择的进化过程，所以大脑的能力可以很好地理解为适应和适应的副产物。对某些人来说，这似乎是显然的，在认知心理学的某些领域，它也确实无可争议。例如，研究视觉的人从来不怀疑眼睛是为了“看”而进化的。但在心理学的其他领域，谈论进化却被认为是不合时宜的、天真的或者在政治上可疑的。威尔逊（E.O.Wilson）的《社会生物学》在1975年出现的时候，作为一部尝试把进化论思想应用于侵略、性和利他主义等领域的现代化著作，却遭到普遍的反对。[36]但在最近10年里，突然涌现出一门新的学科——结合了当代认知科学与进化生物学的进化心理学，它的倡导者是圣塔巴巴拉加利福尼亚大学的一批学者，如柯斯米德斯（Leda Cosmides）和图比（John Tooby）。

在这个框架下提出的许多具体建议现在看来还有争议，但研究纲领本身却得到了越来越多的认同——以至人们可能已经不再把它当作一个分离的心理学领域了。在未来50年，“进化心理学”将成为一个时代的误会，因为那样的标题意味着还存在别的什么心理学领域，不关心选择性优势和适应性设计等方面的考虑。（难道是造物论心理学？）有人不关心大脑如何工作、进化如何发生，同样，也总会有那么一些对进化论不太感兴趣的心理学家。然而，进化论的思想（与神经科学和发展心理学的思想一道）成为心理学研究的证据源泉，不会再有什么争议了。

像这样跟其他学科重新综合，未来50年的心理学将在多方面更加丰富。从传统说，心理学是关于两个群体的研究：大学新生和小白鼠。不过，研究者们至少在不自觉中越来越熟悉多物种的研究了——不是因为天真地相信思维处处相同，而是把它作为认识精神系统演化的一条途径。同样，我们也自然会通过人生的不同转折点、精神疾病的不同形式和不同文化之间的差异，来探索精神生活的基本结构和过程。现在，方

法的多样性在与记忆和感觉相关的领域里是我们研究的标准，但在诸如社会和个性心理学那样更“软”的领域，它也会越来越流行。

我们来看道德思想和行为研究的例子。这在过去曾是心理学探寻的最基本领域。实际上，它也是达尔文与詹姆斯所不同的一点。在《人类的由来》中，达尔文想通过人类智慧的普遍增长来解释人类的道德——智慧的增长使我们超越了我们灵长类祖先的情绪反应，懂得了伦理行为的概念，也就是懂得了公正客观的道德准则。詹姆斯则抱不同的观点，他在《心理学原理》中争辩说，人类本性的独特方面不过是其他动物所缺乏的社会本能（如羞怯和隐私）的总和的结果。达尔文同时代的华莱士（Alfred Russel Wallace）还有另一个观点：他认为人类的利他行为太过神秘，它的存在否定了自然选择理论可以应用于人类精神。总的说来，过去只有哲学家和神学家才研究的利他与道德，如今成了进化心理学的核心问题。

但是学生们在心理学导论课中可能听不到那些东西，却要用几个学时来糊里糊涂地学习老鼠和灵长动物神经系统的大体解剖。我本人的领域是发展心理学，奇怪的是，这个领域里的好教科书为语言学习留下了许多空间（那个领域已经有了几种杂志和专门会议），却没有多少空间来讲道德的发展，如儿童怎样形成一定的是非观点，这些观点又如何对他们的行为产生影响。

这里并没有什么特别的意思，道德发展讲得少，只是因为我们对它的了解太少，不是因为没有兴趣研究它。实际上，道德发展的研究有着重大的实践意义。父母想知道如何把孩子培养成人，有责任心的公民想为年轻人的成长营造良好的环境，如学校教育等。这些都是普通的愿望，尽管人们对什么是好的道德还有着不同的认识。所以，我们想知道针对具体问题的答案：打屁股对孩子好不好？暴力的电脑游戏或影视作品会产生什么后果？父母养大的孩子比单亲家庭的孩子更好吗？托儿所对孩子的气质和情感有什么影响？

尽管我们也许能在报纸或电视新闻中看到一些报道，但所有那些问题的答案我们还是不知道。我们有一些经验性的猜想，不过在根本上我们只能说，好父母培养出来的孩子往往比坏父母养大的孩子更好。不管父母怎么“坏”，不论是虐待、酗酒、精神分裂还是对家长教师协会的漠不关心，这些毛病常常会表现在他们孩子的身上。但我们不知道为什

么。这种现象也许是父母养育的结果。例如，好斗的大人可能培养出更加好斗的小孩。或者说，个性是可以通过遗传转移的，所以父母的好斗与孩子的好斗一定存在什么联系，哪怕他们从来没有相遇过。它也可能是小孩影响大人的结果：好斗的孩子很容易惹看管他的大人发火和动粗。还有很多其他的可能，而且当然很可能存在某种复杂的相互作用。

我写这篇文章的时候，一本专著出版了，它报告了一个大规模的研究：电视节目对570个青少年的影响。他们看电视的习惯从5岁开始；10年后再对他们进行测试，测试指标包括成绩、好斗、吸烟，等等。与多看暴力节目的孩子比较，上学前多看教育节目的孩子长大以后似乎成绩更好，也不大吸烟，很少有好斗的。报告具有重要的政策意义，符合我们的普遍感觉：教育节目好，暴力节目坏。所以，我们的教育节目应该更多，暴力节目应该更少。

不过，在报告的讨论中也隐藏着作者们的另一层意思：他们的发现可能还有别的解释。我们毕竟已经知道，五岁的孩子有的好斗，有的害羞，有的喜欢动物，有的喜欢运动，等等。在研究中，让孩子们选择他们要看的电视节目，可以想象，喜欢书本和知识的孩子更愿意看《芝麻街》和《罗杰斯先生的邻居》[37]，而好斗的孩子可能喜欢看多暴力的节目。所以，那些研究只不过说明，上学前好斗的孩子很可能长成好斗的大人，爱书的孩子很可能长成爱书的大人——这些事实也是我们早就知道的。电视跟它们可能没有一点儿关系。

也许电视确实产生过影响，问题是我们还不知道。我们并不单是需要更多的研究，我们需要的是一个道德发展的理论——一个通过跨学科的研究（包括认知心理学和进化论）来形成的理论。我们需要一个在丰富的语言发展和知觉发展理论基础上的道德发展的理论。只有那样，我们才能切实地谈因果和预防的问题。

我们能在未来50年得到那样的理论吗？直到目前我还是乐观的：心理学将变得更加有趣；人为的学科界线将消失；研究的范围将扩大；等等。这些都是好的方面。不过，至于我们在道德思想或意识等深层问题上能有多少进步，还是悬而未决的问题。另一些人，如乔姆斯基和哲学家迈克金（Colin Mc Ginn），还抱着怀疑的态度。毕竟，我们是人而非天使，有我们能理解的事情，也一定有我们不能理解的事情。也许道德思想或意识的本性根本就超出了我们的理解力，不是因为它们处在什么

特殊神秘的地位，而是因为我们没有理解那些事物的能力，我们大概像是在努力理解微积分的狗。

我们无法知道这些悲观的论调是不是正确，但愿它是错的，我们也只有在这样的希望中迈步向前。不过，直到现在也没有什么进步，这么引人注目的失败该让一部分心理学家谦逊一点吧，特别是他们在制定社会政策的时候。因为没有进步，所以我没有谈心理学在未来50年里可能带来的实际好处——忽略这一点似乎有些奇怪。难道我们不应该希望未来的心理学能治疗精神疾病，清除不幸，摆脱偏见和无知，教我们学会培养一个道德、幸福、独立的孩子，享受其他一切美好的事物吗？这是人们从许多普及读物里得到的印象，许多对自己的能力充满信心，渴望经费资助和政治影响的心理学家，起了推波助澜的作用。

实际上，心理学的现实好处总是有限的。如果不谈某些临床的革新——多数来自生物化学和神经学——往好处说，心理学的一般意义在于教我们如何管理社会、预防犯罪和教育抚养孩子；往坏处说，它们在赶时髦，很危险——例如，在我本人的学科领域里广泛流行着这样的观点：如果不在三岁以前向小孩灌输社会和认知的启蒙知识，就可能毁了孩子的一生。还有，向婴儿演奏莫扎特的音乐有好处，托儿所的日托不好，小孩刚出生的几个小时对母子情结非常重要，等等言论，都是类似的例子。关于这些大众言论，只有一点可以肯定：它们随时都在改变。如果你不喜欢心理学家现在讲的如何培养孩子的那些东西——孩子应该学什么学科，学多少，作息时间怎么安排，等等——你可以等一两年，那时他们会告诉你不同的东西。

乐观地说，如果把在某些领域（如感觉）表现良好的方法和理论，应用于认识薄弱的更“软”的领域（如道德思想），未来50年的心理学会更加成熟。在这个探索的过程中，我们也许能对我们的思维形成足够的认识，对我们的科学方法产生足够的信心，这样，我们才能发现和承认那些问题有多么艰难，还有多少东西需要我们去学习。

布隆（**Paul Bloom**）

布隆（Paul Bloom）是耶鲁大学心理学教授，国际公认的语言和发展心理学专家，与Steven Pinker合写了本领域的开创性论文之一。他是耶鲁最年轻的正教授，发表了50多篇（章）心理学、语

言学、认知科学和神经学的文章。他还写了《儿童如何学会生词的意义》（How Children Learn the Meanings of Words ），即将出版《肉体与心灵》（Bodies and Souls ）。

捉摸不定的心理学

G.米勒

Geoffrey Miller

当我们把自己想象为感情动物时，有太多的人成了“小气的”唯物主义者。这种狭隘的唯物主义者抱着这样的观点：如果主观经验尚不能与大脑的某个特殊区域或神经传递介质或基因联系起来，那经验就可能不是真实的。他们认为，当我们发现了与某种疼痛相关的大脑区域时，那疼痛才真正具有科学的意义——但是，假如我们还没有发现与某些感觉关联的大脑区域，如两性的嫉妒、存在性恐惧等[38]，那么这些情绪就可能不是“真实”的，应该用怀疑的眼光来看它们。同样，假如我们在精神分裂症患者身上发现了神经传递介质缺陷，那它就是真正的病态；而如果我们在坏脾气的人那里没有发现类似的缺陷，那么坏脾气也许就不是真的病态，只是个性的缺点或坏习惯而已。

狭隘的唯物主义者不仅不相信他们自己的意识，也不相信他们的唯物论学说在未来的进步。结果，他们盲目迷信神经科学，寻求它做一切主观事物的根据。因为神经科学至今还在萌芽中，过分依赖它那有限的作用，只能产生对人类本性的幼稚认识，仿佛人就跟卡通画一样简单，几根线条、几块色彩就勾勒出来了。

像达尔文和詹姆斯那样的“大方的”唯物主义者则抱着截然不同的观点，他们相信一切捉摸不定的思想和感觉的东西都基于大脑的行为。因为同样相信科学的唯物主义和人类意识，他们能慷慨地把丰富的主观经验归结到人类和人类复杂的大脑。我想，在未来50年，我们自己的唯物主义也会像达尔文和詹姆斯那样变得更加慷慨大方。当神经科学揭示人类意识的更多微妙，人们会更容易接受和认识那些微妙，给我们和我们的社会带来好处。狭隘的唯物主义者容易使人自私和自大，因为神经科

学似乎只看重我们与动物共有的更原始的能力和情感。我们更加进化的有别于动物的人的能力，如创造、善良、幽默和想象，还在脑图像实验室里演习。大方的唯物主义者则可能使我们更多情、更谦逊；我们会发现其他思想和头脑也有着我们那样丰富的主观经验。

为了这个修正的更加慈善的人性观，关键的一点是发展新的技术，揭示大脑的神经活动和基因激发模式。不论当我们从事没有特殊感觉的简单的认知活动，还是参与最能激起人类情感的复杂的社会活动，它都能告诉我们脑海里到底发生着什么事情。当新技术客观证实了我们纷纭复杂的思想、感觉和社会关系，不论每个人的心理复杂性，还是人与人之间的差异，都将在科学的进步中形成更加精细的认识。

一个世纪以前，我们不得不靠亨利·詹姆斯的小说，通过精确的细节和富丽的色彩来刻画人的意识。[39] 将来，我们不能指望大众文化来做那种事情——Viacom [40] 和迪斯尼（Disney）看不出那有什么好处。不过，我们也许能通过科学来填补这个空白。

一段短小的历史可以帮助我们正确认识狭隘的唯物主义。在19世纪，心理学是西方男性资产阶级知识分子从事的学问，也是为他们的利益服务的学问。在这样的家长制心理学中，几乎不关心妇女、儿童、非知识分子、其他文化背景的人和其他有意识的动物。但是，这种严格封闭的思想方法却有着被人忽略的优点：西方资产阶级知识分子共有的心理特点使他们能发展一些精密的方法来表现和传达他们丰富而捉摸不定的内心生活。资产阶级文化与资产阶级心理学之间存在着一条封闭的正反馈环，这典型表现在詹姆斯兄弟威廉和亨利的往来书信中。他们在内省 [41] 中热烈交流，形成了刻画意识的超凡本领。更一般地说，19世纪后期欧洲文化的演进就表现在达尔文、高尔顿（Francis Galton，达尔文的侄儿）、弗洛伊德、哲学家勃伦塔诺（Franz Brentano）、美国心理学家鲍德温（James Mark Baldwin）和迈克道戈尔（William McDougall）和德国实验心理学家冯特（Wilhelm Wundt）等人的心理学理论化过程中。他们敢于猜想情绪、美感、爱情、家庭生活甚至变异的意识状态。

随着大众文化的民主化和还原论与实证论在科学中的兴起，这种风格的心理学在20世纪都烟消云散了。心理学拓展了研究的对象——让妇女、儿童、工人阶级、非西方人和灵长类动物进入它的领域——但是紧缩了研究的问题。同时，随着电影、广播、电视等大众媒体的兴起，文

化产业有了更多的顾客，也失去了过去占主流地位的小说和戏剧的亲和。在这些媒体的观众看来，人性的文化形象多了僵化而少了敏感。与此相应的是，心理学的内容也在总体上大为简化了。20世纪20年代华生[42] 和50年代斯金纳的行为主义把主观经验看作错觉，把学习作为行为的基础，把几乎所有其他的事物——思想、感觉和社会的相互作用——都从心理学赶了出去。

随着60年代的认知革命，计算取代了学习而成为心理学最流行的隐喻。一些心理学家从而解放出来发展知觉、推理和谈话的计算机模型。不过，某些模型并没有扩张多少学科的主题。严肃的心理学家仍然忌讳写真正的社会、性和家庭关系或者多变的意识状态，如浪漫的爱情、父母的骄傲、职业的嫉妒。实证论、经验论和还原论把证明的负担转移到了那些想考察人类意识的人身上；任何不能在实验室证明的思维状态都被认为是不真实的。由于几乎所有关于人类有意识的内心活动的东西都不可能这样袒露出来，在那时可能的实验条件下，大多数主观的人类活动都被排斥在心理学的大门外了。

总的说来，西方文化和西方心理学在20世纪都美国化了。它们的领域更宽广，却不那么精密；方法更客观，结果却不那么准确；政治上更加进步，却不那么有人情味。它们也关心个体，不过是被剥夺了社会、性和家庭关系的分化了的陌生对象。最后，它们能更加有效地描述和探索对广告和宣传的简单的下意识反应，却不大接受涉及模糊、想象、同情、道德判断或美学抉择的任何有意识的状态。心理学认定它不得不在分析行为与理解意识之间、在经验的重要性与主观的精确性之间做出选择。

我想未来50年将证明这种主张是错误的。新技术有能力证实更多的人类的主观经验。只要科学家的工作是研究在有意义的社会条件下的真实思想和意识，就可能产生更丰富也更精确的人性模型。

例如，脑图像方法能说明当我们在实验室进行一定的精神活动时，哪些脑区域在活动。直到目前，多数这类实验都是根据标准的知觉心理学或认知心理学来做的，对参与者没有内在的意义。读大学的时候，我也做过这样的脑图像实验。在哥伦比亚医学院，我被绑在床上，头连着20个盖革计数器，呼吸着氧和放射性氙的混合气体，花6个小时在简单的几何图形间进行选择。当然，实验有助于我们识别与图形识别有关的

大脑区域。但我在考虑那些图形的时候，还有许多跟图形识别不相干的思想和感觉：担心我和女朋友刚吵了架，回想帕索里尼 [43] 的电影情节，想着里根总统的老态龙钟。从研究者的观点看，这些偶然的念头都是“噪声”，只要实验里的这些不相干的思维状态具有足够的多样性，研究者可以放心噪声最终是能被消除的。

然而，实验还有一些特别的事情，似乎说明参与者的思想和感觉要比实际的简单得多。我们应该承认，目前的脑图像技术还不能很好满足我们研究现实社会状况下流动而复杂的思想和感觉——我们应该努力发展那样的技术。在一定程度上，空间和时间分辨率更高的仪器更能有所帮助，假如能精确测量在1毫秒×1毫秒基底上有多少个立方毫米的大脑区域在活动，我们就能研究更微妙的心理学过程。另外，研究作风也需要改变。我们可以继续发扬19世纪勃伦塔诺和詹姆斯实践的传统的“内省”方法。在内省中，心理学家自己就是他最好的研究对象。我们可以让自己坐在脑图像仪下，系统考察我们的思维状态，看看什么在活动。然而这跟闭门自省有同样的局限：我们不可能在关注大脑状态的经验数据的同时，还继续进行各种带着强烈主观色彩的现实的社会活动——如谈话、调情、讨价还价、争辩、哺乳，等等。为了做到这一点，我们需要轻便、结实而又不显眼的脑图像系统，只有那样我们才能将大脑全部的真实的能力展现出来。

确认人类意识复杂性的另一关键技术是把基因的表达模式画出来。每个大脑细胞都有一组完整的基因，但在任何给定的时刻只有某些基因能表现——就是说，只有某些基因能转录到RNA，然后转录到蛋白质。而且，不同的大脑区域有着不同的基因表达模式，这些模式随时间而改变，不仅在从胚胎到成熟的演化过程中改变，也随日新月异的情况而改变。在社会环境、神经心理学、基因表达和行为之间，存在一个反馈的环。当我们恋爱，当朋友去世，当我们得到升迁，我们大脑的基因表达模式无疑会改变。几乎每一个持续时间超过几个小时的思想过程都包含着基因表达的改变，科学家刚开始跟踪这些变化。

一旦技术进步，我们能实时跟踪基因表达，一个新的复杂的心理学世界将出现在眼前。我们将看到，现代的社会处境可能产生相应的遗传进化影响下的行为能力。我们还将摆脱所谓环境与遗传“混淆不清”的胡言乱语，更清楚地看到一定的处境、思想和感觉如何激发一定基因，一定的基因又如何发生反作用。当我们在我们整个大脑里找到进化的遗传

印迹，就不会再有人指责进化心理学不过是一堆“如是我闻”的故事了。

如果我们有勇气以同情的态度来运用脑图像和基因表达图的新发展，这些新的技术将说明更广泛的人类经验。如果能在意识状态下——现在看来那是短暂而奇异的状态——发现真正的神经和遗传的印迹，那么我们会更严肃地把那些状态作为人性的一个普遍的组成部分。我们本不该需要这样实在的证明，但我们现在确实需要：似乎存在那么一种天生的倾向，个人自己的精神生活仿佛比其他任何人的更复杂，更有意义，也更正确。19世纪的内省家们把精力集中在自己身上，忘了别人，描绘了一个丰富的内心世界；20世纪的行为主义者们忘了自己，把心用在别人身上，在学习和计算的基础上描绘了一幅粗野的人类本性图像。21世纪的心理学家们，将从人类意识那倏忽无常和捉摸不定的表现中发现神经和遗传的信号，从而打破自我与他人、主观与客观之间的分野。结果应该是一个更有人情味、更加包容的人的科学。我的希望是，2050年的大学新生在上心理学入门课的时候，会惊叹“啊哈！原来我们就是这样从y感觉x的呀！”而不会像我们今天到处听到的那样抱怨，“说到底，这跟我们的现实生活有什么关系呢？”

米勒（**Geoffrey Miller**）

米勒（Geoffrey Miller）是新墨西哥大学进化心理学教授，《杂交的思想：性选择如何形成人性的进化》（The Mating Mind: How Sexual ChoiceShapedtheEvolutionofHumanNature）一书的作者。

幸福的未来

M.西克斯詹特米哈依

Mihaly Csikszentmihalyi

有一个问题将在未来50年成为焦点：我们应该如何发挥控制人类基因组成的能力？过去，我们的先人用粗糙的遗传选择方法来决定哪些小孩活到哺育后代的年龄。现在我们通过科学的帮助也掌握了实现相同目

标的那种可疑的本领。

远在人们猜想存在基因之前，农民就发现母体的特征传给了后代，于是他们能通过有选择地培育最好的品种来提高南瓜的产量或猪的重量。那时很容易把这种原理用于人类。柏拉图在《理想国》第五卷中用了很大的篇幅讨论如何用培育猎犬的方法来为他理想的国家培养统治者。例如，他在第459章写道：

> 最好的男人应该尽多地与最好的女人结合，而最坏的要尽少地与最坏的结合……假如要保持最好的品质，他们应该养育好的后代，而不应该养育坏的后代。这些过程应该只让统治者知道，否则会有激起……叛乱的危险。

在更前面的第三卷第415章他写道，“神晓谕统治者的第一原则是……他们必须护卫种的纯粹，没有比这更需要他们认真维护的东西，或者说，没有什么东西更需要他们做那样好的护卫者。”实际上，我们知道的任何社会都有过这样的习俗，今天也许可以为它们贴上“优生学”或“遗传工程”的标签。这些行为常常能在与生物学毫无关系的东西那里——如宗教或风俗——找到正当的理由。不过，它们的实行大概是因为人们认为它们有助于种族的延续。我们应该记得，人人都有生育后代的权利是近代才有的思想；过去的社会则把生育的特权都给了那些可能产生优秀子孙的人们。

积极的习俗促进了有令人羡慕的外表（包括健康、强壮和美丽）和成功的事业（如财富和权利）的个体相结合。门户不等的人，可以通过各种方法来达到结合：结婚前要彩礼和嫁妆几乎是普遍存在的习俗，它保证未来的父母能有足够的财富和家庭支持来抚育他们的孩子，不让他成为社会的负担。

消极的习俗阻碍那些具有不受一定社会欢迎的特征的人生育后代。其中一些几乎是顺其自然的。例如，穷人和病人不愿结婚生子。但另一些却是更主动的行为，如阉割或把婴儿杀死。本来为着其他不同目的的文化习俗，却常常可能产生实在的优生学影响。例如，俄罗斯东正教有一个仪式：把新生婴儿赤裸裸地浸在冷水中，让他们接受圣灵的恩惠，使灵魂免受永恒的诅咒。这个习俗带来一个意外的后果：身体不好的婴儿会死在洗礼中，于是他们的基因就从基因库中清除了。我们只能想

象，这些仪式能流传下来，是不是主要因为它们的虔诚或它们产生的遗传优势能带来内心的安宁？也许，它们是很多因素决定的，宗教和遗传的优势，相对于当时文化的其他习俗来说，更容许它们的存在。

那些习俗多半漫无目的，在认识不同的特征如何从一代传到下一代时，它们没有任何基础。不过，这种状况在即将到来的几十年里肯定会发生根本的改变。当前人类科学中最有活力的两个分支：一个是行为遗传学，它试图确定像精神分裂、离婚倾向、政治信仰甚至幸福生活等行为和品性有多大的遗传能力；另一个是进化心理学，它寻求这些特征赖以选择和遗传的机制。两条路线都认定先天遗传与后天培养在形成我们的行为、思想和感情的过程中都发挥着潜在的作用。不过跟20世纪的学习效应（learning bias）不同的是，它们更看重先天的作用。

随着遗传学的进步，这股潮流在未来的半个世纪里一定会更加汹涌。尽管几乎没有什么重要的特征是靠一个或几个基因的作用决定的，还是有些遗传工程学家相信，“设计师的婴儿”的时代已经到来。即使他们过分乐观了，我们也不能愚蠢地对可能马上就要面临的抉择视而不见。我和我的同事们在最近的一个研究中访问过100位前沿的人类遗传学家，有趣的是，他们不在乎研究中的那些有着更多争议的东西，多数认为那跟优生学之类的事情没有任何关系。他们不相信人类能克隆自己，不相信自己的发现可能被滥用。他们几乎一致地宣称，关于基因工程的潜在应用，他们没有特别的知识，也不承担任何责任；他们坚持认为那属于政治抉择，应该由整个社会来决定——尽管“社会”还缺乏做出合理抉择的专门认识。这种状况跟半个多世纪以前原子物理学的情形没有什么不同——即使玻尔（Niels Bohr）那样的大思想家，在40年代还认为核裂变实验不可能有实际应用。

但是，不论是否情愿，我们需要马上做出抉择，而那些抉择将决定我们的未来。例如，我们设想可以很快地极大提高普通智力因子g——它是教育体系高度重视的语言和数学技能的基础，在其他生命领域也起着重要作用——这个主意好吗？[44] 几个评论家指出，社会已经被知识分化到了令人厌恶的程度。在不远的过去，勤劳、诚实、友好和善良的人，不必“才思敏捷”，也被认为是成功的；而今天，抽象的思维技能正在成为成就任何功名的先决条件。假如我们发现了强化这种特质的遗传方法，人类的智力将指数式地增强。如果“超聪明”的人和普通人之间距离拉大了，在他们的经济和政治力量之间也会出现巨大的鸿沟。基于知

识的“同族婚配”已经出现了，还将更加流行。一个智商超过200的人做梦也不会想跟一个智商不足150的人结婚。假如这种工程影响了族谱，这些分化将自动传给下一代。

不过我们还可以想象，假如我们发现了增强每个人的智力的方法——那当然是不大可能发生的平等主义的故事——从而可以提高整个人类的智力底线，结果会怎样呢？那想法好吗？答案是我们不知道。大多数生物学和心理学功能只能起“小剂量”的作用，一旦过分就变得危险了。如亚里士多德指出的，美德到了极端也成为缺点：如勇敢变成鲁莽，谨慎变成多疑。天才与疯狂的模糊关系意味着过分的聪明也可能成为自己的障碍——例如，过于敏感容易产生焦虑和抑郁。或者，当理性的力量扩张到了与兰德 [45] 的自我中心态度相结合的地步，就可能产生一个比我们现在更残酷、更无情的种族。

更基本的问题是，如果有了那些方法，那么在改动人类基因组的时候，我们的目标是同一还是多样？同一肯定会有很大的压力：每个人都想自己的孩子聪明，漂亮（公认的美的意义），雄心勃勃，事业成功。多样性却冒着风险。谁愿意把赌注下在未知的和未经证实的事物上呢？不过，生物学家威尔逊（E.O.Wilson）赞同生物多样性的观点仍然适用于心理学特征；越来越趋向同一的种族，不仅危害着人类的敏感性，从严格的生存观点来看也有着潜在的危险。因为未来很难预料，所以最好的策略是拥有多样的潜在能力，这样，我们才可能对突然出现的新环境产生适应的反应，而不致将自己锁定在当前条件下的所谓最好的模式。

假如人类遗传工程能在市场驱动下进行（而不受一个中央计算机的指挥——它将决定下一代的社会需要多少战士、多少工人、多少懒汉），那么最大的需求很可能是哺育幸福的孩子。如果问父母对孩子最大的希望是什么，一般的回答是，他们希望孩子受到良好的教育，找到好的工作；而最重要的是，不论他们如何选择自己的生活道路，都应该是幸福的。当代的父母似乎赞同亚里士多德，他们懂得，如果说其他好东西是通向某个目标的方式，那么幸福本身就是那个目标：它是我们希望通过教育、金钱、美丽和智慧来获得的东西。假如可能通过遗传工程来产生幸福，那将成为父母的第一选择。

行为遗传学家研究了同卵和异卵的双生子，让他们一起成长，然后彼此分开。结果表明，至少50%的幸福是遗传下来的。我们有理由对这

些研究如何度量“幸福”提出疑问，不过有一个事实看来已经很清楚了：存在一个幸福的基点，它在人与人之间有所不同，而相对说来不受外在环境波动的影响。当然，一个人群的一般幸福水平也受经济条件（在某种程度上，钱与幸福是相关的，但是超过了一定的收入底线——例如葡萄牙和韩国的平均水平——更多的收入并不能带来更多的幸福）、政治地位和许多其他外在因素的影响。但不管怎么说，个人的遗传仍然起着很重要的作用。

所以我们设想，在未来的几十年，有可能通过遗传工程来增强我们孩子的幸福。我们会利用这个机会为他们做好事吗？社会和整个人类会从这个抉择获得好处吗？这些问题的答案会怎样呢？我们可以在这里先回顾一下我们仅有的一点关于幸福的认识。

首先似乎可以肯定一点：人们对幸福的自我感受，是幸福的一个十分恰当的度量。它还密切联系着家庭和朋友的感受，联系着病理学和相关的行为——总之，自以为幸福的人，看起来和做起来都应该像是我们认为的幸福的样子。他们喜欢交朋友，有稳定的社会关系，过着健康和富裕的生活。这些都是没有疑问的。

但是，也可能存在一些有趣的然而消极的方面。例如，幸福的一个最为流行的定义说，幸福是一种不再有任何其他欲望的生活状态。幸福的人不太在乎物质财富的多少，不受广告宣传的影响，不为功名利禄所驱使。为什么会那样呢？他们已经幸福了，不是吗？幸福者的社会观大概足以彻底粉碎我们建筑在不断增长的消费和永不满足的欲望基础上的生产体系。

理论心理学能为这些迫在眉睫的抉择提供什么答案吗？大约20年前，这门学科还几乎不谈幸福。对严密的科学研究来说，幸福的问题太“软”了。为了走出困境，心理学将不得不再回到它原来的对象：心智——不是飘忽神秘的灵魂似的东西，而是当我们的注意力转向分析、综合和回应外在刺激或内心状态（也就是思想和感情）时，从我们的意识产生的一系列非常具体的现象。意识流在多数科学家（包括心理学家）看来太过主观了，不适合严格的科学研究，但事实上那是我们能获得的最客观的数据。科学事实和基于它们的知识是别人说的，我凭信仰可以愉快地接受；而意识的活动，如恐惧、欢乐、愤怒和希望——我能亲身经历，其实在性是不容争辩的。

对我来说，我决定发展一门系统的现象学，它将回答以下的问题：人们的思想、感觉、目标和行为，在普通的一天是如何波动的？在人的一生又是如何波动的？意识流的这些不同部分是如何相互联系的？在日常生活中人们何时感觉幸福？每一个这样的问题都可以引出几十个更进一步的问题，还包括，年龄、性别、种族等诸如此类的差别如何影响意识，某个时刻测度的模式与一年以后的模式有什么联系。这些事情中，我们所知道的是，怀着明确目标紧张投入挑战活动的人比那些过着轻松安逸生活的人更幸福。一个人为自己工作越少，社会关系和承担的义务越多，他就越感觉幸福。

还需要认识到，意识有其自身的具体的实在性，当我们拿适用于不那么复杂的系统的方法来分析它时，一开始就把它破坏了。它是一个开放系统，状态不停地随时间而变化。例如，即使你已经掌握了60秒钟之前我大脑的一切信息，如化学、遗传背景、过去学过的东西，等等，也不可能根据我1分钟前的思想来准确预言我这时在想什么。在那两个时刻之间发生的事情，不论听到的、看到的或感觉到的，都可能在1分钟内进入我的意识，以全新的不可预料的过程形成我的思想和感觉。

在创造性活动中我们可以清楚地看到意识的不确定性。普遍认为，一首诗（或一支奏鸣曲、一幅画、一个科学理论）的基本要素可以从诗人的意识中找回来——只要我们完全把握了他的意识的所有信息。就是说，我们跟胚胎发育的小人理论[46]是遥相呼应的，相信创造性工作包含在创造者的身上——尽管只表现为某种微观的或密码的形式。诗人可以从一个字或一个短语开始——原来没有什么意思的普通字眼在那个特别的时刻似乎令他特别感兴趣。如果我们知道诗人刚才的意识和感觉，也许就能解释为什么那个字眼突然有意思了。但是接下来发生的事情就不知道了：那个字眼可能引出无法预料的思想或联想，而那些引出的东西又打开了新的思想和感觉的方向，引出更多的文字。这样，一个意境的圆圈在一个突现的、自组织的系统中逐渐展开来——它的基础还是诗人开头的意识，但已经不可能再还原了。

我们用不着借创造性活动来说明这个过程。让我们来看一个更普通的例子，看父母对他们的新生宝贝有什么反应。遗传和进化心理学能告诉我们，父母如何与他们的孩子相亲，他们为什么会那么亲密。养育孩子是最古老的人类经验之一，自从有了人类以来，它就是每一代人的经验。不过，一个人即使知道所有关于生孩子的事，第一眼看到自己的孩

子仍然是独一无二的经历，不可能为它做好充分的准备。捉摸不定的地方在于一个人如何认识自己的配偶、经济状况和生活里的一切——更不用说宝宝的模样、身体和行为——所有这些因素都可能与宝宝的出生这件大事有意义地结合在一起。你可能想通过尽可能多地了解父母来猜测那些因素会产生怎样的联合作用，不过那种猜想是不会准确的，因为影响父母意识的还有太多外在的变化的因素。

假如心理学要把意识流纳入它的领地，它可能为我们选择我们向往的光明的未来提供需要的知识。知识每多一点进步，我们就多一分责任。过去，我们仿佛坐在进化的四轮马车上的悠闲过客。今天，进化更像飞过太空的火箭，我们不再是乘客，而是宇航员。我们会创造出什么样的人来呢？是像我们的机器和计算机那样的血肉躯体，还是朝着前所未有的方向快乐进化的有着向宇宙开放的意识的生命？

心理学开始显现出沿着后一个方向前进的迹象。在美国内外的众多研究中心，严肃的学者们正在研究智慧、生命目的、内心动机、灵性等课题——在几十年前，它们还都是门外的话题。美国心理学会主席塞里曼（Martin E.P.Seligman）在最近的任职期间掀起了专业领域的“积极心理学”运动，超越了传统的治疗精神痛苦的目标。今天，运动的成果之一是拓展了一系列普遍存在于不同时代和文化的“力量”，如智慧、勇敢、坚韧和诚实。接下来的一步是把如何运用这些力量的知识集合起来。最后，这些知识应该渗透整个专业，使它与治疗和预防实践有同等的重要性。我们需要那样的科学来迎接未来50年的挑战。

西克斯詹特米哈依（**Mihaly Csikszentmihalyi**）

西克斯詹特米哈依（Mihaly Csikszentmihalyi）是出生在匈牙利的博学者，曾为芝加哥大学心理学系主任，现在是加利福尼亚克莱尔蒙特大学戴维逊管理教授。他在最佳体验心理学的研究和理论已为国家领导人克林顿和布莱尔以及许多世界大公司总裁们付诸实践了。他的著作包括畅销书《流：最佳体验的心理学》（Flow: The Psychology of Optimal Experience ）、《进化的自我：第三个千年的心理学》（The Evolving Self: Psychology for the Third Millennium ）、《创造力》（Creativity）和《发现流》（Finding Flow ）。[47]

50年后我们还会忧伤吗

R. M.萨波尔斯基

Robert M. Sapolsky

当我们走上新世纪的征程，忍不住想做两件事情：一件是总结——把荣誉献给刚刚过去的一个世纪的最重大的事件或成果；另一件是展望——对我的某些同行来说，就是选出某个20世纪的疾病，努力去认识它未来的行迹。

考虑选择什么病的时候，也许有人会关注1900年以来被攻克了的那些疾病。在那个范围里，逻辑的选择是天花——它的消灭是医学的一个伟大胜利——但从感情上讲，人们可能更愿意选择小儿麻痹症，尽管不幸的是它还在很多第三世界地区肆虐。在西方，人们现在还记得小儿麻痹症肆虐的20世纪初的铁肺恐惧。[48] 许多小儿麻痹症的幸存者在50岁以后还受着后遗症的折磨，那是最后的痛苦，衰弱但还没有崩溃的神经肌肉系统已经老化了几十年，肌肉变得衰弱和萎缩。另外，萨宾（Sabin）与索尔克（Jonas Salk）的争论也还有耐人寻味的地方。[49] 当然，我们的选择也许应该落实在那些在整个历史进程中几乎以相同速度蔓延的疾病——现代科学对它无能为力，不也是很好的新闻卖点吗？在这种情形，可以考虑疟疾。20世纪还出现过一些令人谈之色变的疾病——如艾滋病以及跟它同样可怕的癌症、心脏病、糖尿病和阿尔茨海默症。

但是，如果要选择一种疾病，既是毁灭性的，又对现代医药具有惊人的抵抗力，而且已经流行起来，那么我想应该是“大抑郁”（Major depression）[50] 。

我说的“大抑郁”，不是令我们日夜担惊受怕的失败和挫折，不管它是什么，我们最终会发现它不是世界末日。大抑郁令患者度日如年；他们陷入绝望的深渊，不能工作，不能社交，不能爱，不能睡，不能吃。他们甚至活不下去了，几乎一半的人在不同的时候想过自杀。抑郁是标准的现代精神疾病：它是涉及基因、神经化学、荷尔蒙等诸多方面的一

种导致精神“病”的生物学障碍，这种障碍对环境特别敏感，患者容易产生绝望和无助的感觉。

大抑郁的流行触目惊心，大约15%的发达地区的人在一生的不同时期经受过抑郁的折磨。抑郁正变得更加普遍：最近50年，西方国家的抑郁人数在持续增加。也许有人怀疑这个事实潜藏着虚假——因为今天的抑郁者比过去更愿意寻医问药，而保健专家似乎也比50年代的医生更有希望诊断抑郁——为了澄清这样的混乱，这些研究是精神病学中前所未有的最严格的流行病学研究，而且是在周密控制下进行的。抑郁人数确实在不断地增加着。

那么，大抑郁在未来50年的命运会怎样呢？不幸的是，我怀疑这个医学灾难不会消失，而且还可能更加猖獗。

凭什么得出这个结论呢？首先，关键的一点是，我们需要懂得紧张与抑郁之间的联系，还有我们的生活怎么会越来越紧张。我们（和我们的身体）很可能不堪忍受某些外来的挑战——假如我们缺乏控制它的意识；或者不能预感它什么时候到来，不知道它会变得多么恶劣；或者没有社会的支持，无法排解由它引起的失落。抑郁的一个非常有用的模型是心理学家塞里曼（Martin Seligman）和他在宾夕法尼亚大学的同事在20世纪70年代发展起来的。模型叫“习得性无助感（Learned Helplessness）”，就建立在刚才说的那些因子上。面对心理学的紧张刺激时，我们多数人能把它放在一定的背景下，限制在一定的范围内，知道它不能代表整个世界。就是说，我们能应付它。当我们失去那个界限，而且将它无限扩大，抑郁就产生了，也产生了扭曲的结果：“不但这件事情可怕，我控制不了，所有的事情都可怕，我的生活中没有一件我能把握的事情。”抑郁患者就在那样的经历中感觉无助。强烈的紧张可能把所有的人都推向那种幻觉，而抑郁的生物学风险更大的人会更加脆弱。从一定的生物学水平说，大抑郁是紧张后的失调，它屈从于彻底的无助感，并最终走向绝望。

为什么越来越多的人变得抑郁？我为什么认为还会继续多下去？运气好的话，应该有很多东西使抑郁的流行慢下来。例如，50年后，小女孩很可能比她们的前辈更能决定长大以后做什么样的人——神经外科医生、公司总裁或者足球明星。制度下的种族隔离、犹太人配额[51]、“不招爱尔兰人”的牌子[52]，都将成为模糊的历史。许多长期遭受歧视的人

们，将来不会再经历那样无助的煎熬了。而在某些方面，我们的地球也许真的会变得不那么严酷：我们可以乐观地预言，越来越多的人至少表面上将过着自治的生活。奴役、强奸和寡妇自焚 [53] ，同样也将在世界范围内减少——当然，也许那是我们对人类的过高期许。

另外，科学发现了越来越巧妙的同大抑郁搏斗的办法。我们认识了一些神经传递介质——大脑里的化学信使——它们可能在抑郁中出现反常。最显著的例子是5-羟色胺，它有许许多多的功能，包括影响很可能与抑郁相关的情绪调节。目前我们最多可以猜测，在抑郁中，要么是作为信使的5—羟色胺太少，要么是目标神经元对5—羟色胺传达的信息太不敏感。这个观点的最有力证据是有名的抗抑郁药物氟西汀的作用。它恰当而有选择地增加在神经元之间传递信号的5—羟色胺。下一代氟西汀就在眼前了，它的作用会更快、更强，而副作用比以前更少，如偶尔出现的男性患者的性功能紊乱、记忆力下降、精力不集中等。我们还认识了紧张、失望和某些荷尔蒙（在这种状态下还是一个谜）如何能产生抑郁的神经化学变化。根据这些发现，我们正在荷尔蒙水平上寻求治疗抑郁的新途径。

我们对抑郁的认识还在一点点地增加。在许多长期的抑郁病人的大脑中，似乎有些区域异常地小——特别是海马状突起，它对一定类型的记忆的形成起着关键作用。有长期严重抑郁病史的人都存在着这种可能与海马状突起萎缩相关的记忆缺陷。我们也在学习遗传对抑郁易感性的影响——例如，当5-羟色胺在大脑内传递信号时或者在紧张荷尔蒙的合成中，遗传因子起着什么不同的作用。

这些发现一定能开辟新的治疗前景，而且已经在我们的工作中初露端倪了。当我们逐渐认识了使人与人成为不同个体的生物学时，也许会出现最激动人心的进步。跟在那些认识的后面，我们将更好地认识抑郁的个体诱发因子。

有了我们的科学进步，有了减轻我们无助感觉的因子，我为什么还说我们会更加忧伤呢？主要是因为我想着我们今天的文明里还有许多激发抑郁的东西。我能想象，一定有人会愤怒地批判我正在把人引向邪路：“你要抱怨当今世事艰难吗？那么杜奇曼夫人（Barbara Tuchman）的欧洲呢？大抑郁呢？听说过第二次世界大战吗？”

据我通过杜奇曼夫人那“遥远的镜子”对14世纪生活的认识，我完全有理由相信，那时的每一个人都可能很抑郁，中世纪西方人的精神跟我们今天是迥然不同的。[54] 不过现在，近50年的不断挑战特别可能产生抑郁——主要是因为，离开了社会的支持，生活里越来越多地发生着令人抑郁的事情。过去给我们带来安慰的东西，在即将到来的年代一定会逐渐变得软弱无力。“家庭”将不得不面对有增无减的离婚率；稳定的社会联系将不得不屈从于我们珍视的活动和隐私的自由。今天，在我们的文化里，那种一生都生活在亲戚朋友包围的小村庄里的人已经十分罕见了。

另外，我们的技术似乎不会帮我们缓解紧张，尽管（或者也许可以说因为）我们希望那样。我们会继续推出能节约时间的发明，然后像过去一样，又增添几分多做些事情的希望。我们将拥有更多的物质享受，但我们会重新定位对权利的基本认识。我们有数不清的小玩意儿，有享不尽的闲适和安逸，但是，当我们为早餐的米粥，为整容手术，为新款汽车或者新式婚礼犹豫不决，不知哪样更令我们快乐时，才发现这些选择往往都是徒劳的。

对总体的无助感觉来说，我们的许多紧张性刺激都有特定的来源。虽然在我们可能更加文明的未来，野蛮会受更多的制约，但那些破坏规矩的人也可能有着更强的技术，那不是挥舞棍棒的城市流氓，而是兵器库或者“民兵”。不论我们的全球媒体村会从它今天仅有的500个有线频道变成什么，我们都能通过它的视野亲临每一个惊心动魄的恐怖场景：邻近城市的枪击、另一个大陆的屠杀、凄惨的垂死儿童、退化的生态系统……

不过，抑郁的人可能增加，还有一个关键的统计学的原因。当我们努力从20世纪末的回光挣脱出来——我们在那个世纪里看到了种族可能净化、校园可能成为屠场、第一家庭可能肮脏龌龊 [55] ——我们的孩子却正坐在我们的身边。看看两个重要的事实吧：抑郁越来越多地在青年和少年中发生；小时候大的紧张刺激肯定增加成年时患抑郁的风险。儿童时代本该学习控制外在环境的能力和极限，学习什么是可以依赖的快乐源泉。而我们却在孩子更小的时候就让他们知道了邪恶的秘密——世界到处是痛苦和悲哀，而我们却无可奈何。面对这样的现实，没有哪个小孩能像大人那样形成抵御的防线。有的人可能会超然物外，好好生活是对它最好的回应，这种态度理应成为一种风尚；另一种态度则更多地

在那些顾影自怜的人中间流行，他们更脆弱，更容易绝望。绝望的种子已经在下一代那里萌芽了。

那么，化学药物能带来好生活吗？——我们未来就要靠那些药丸来摆脱痛苦，换取安宁。目前所有的抗抑郁药物都不是很有效的。许多抑郁症患者因为不能忍受那些副作用，不得不把药停了。其他患者，有的只有在经过多年的不同药物实验后才能得到一定的缓和；还有很多连缓和也得不到。难道我们没有成功的可能，“让抑郁滚开”的药物最终将把那种疾病赶进档案馆里？显然，这样的事情不会发生，科学家也知道。写这篇文章的时候，我想找一个医药公司的市场副理，希望他能给我一张这样的药品广告：“抑郁吗？太简单了——这20年就靠它了。”可惜没有这样的人。即使对控制抑郁抱乐观态度的人，也不抱那样的希望。

这一点儿也不奇怪。医学可以通过多个途径取得进步——我们曾经抽干沼泽的水来消灭蚊子；在非洲的某个小村庄追溯最后一个天花的病例； [56] 我们也认识了细胞如何产生癌变，个别马基雅维里式的病毒如何破坏原以为能破坏它的免疫系统。 [57] 但在抑郁的医学问题上，那些方法都显得软弱无力。从来也不会有什么关乎生命荣枯的疫苗。到头来，一定有人要问的问题和研究抑郁的科学家、临床医生甚至进化心理学家也一定要面对的问题，不是为什么我们将有那么多人向抑郁屈服，而是我们多数人应该怎样控制和避免它。

萨波尔斯基（**Robert M.Sapolsky**）

萨波尔斯基（Robert M.Sapolsky）是斯坦福大学生物学教授和斯坦福医学院神经学教授。他还是肯尼亚国家博物馆研究员。他原来关于紧张和神经疾病的研究是在实验室进行的，在23年里他每年都去东非塞伦盖提（Serengeti）大草原研究狒狒种群和这些动物的个性与紧张性疾病之间的关系。他最近的那本《灵长动物实录》（A Primate’s Memoir ）就是多年非洲生活的结果。他还写过一本《紧张、老化的大脑和神经元死亡的机制》（Stress, the Aging Brain, and the Mechanisms of Neuron Death ），两本为普通读者写的《睾丸激素的困惑及其他人类困境的生物学随笔》（The Trouble with Testosterone an Other Essays on the Biology of the Human Predicament ）和《斑马为什么不患溃疡：紧张、紧张性疾病应对指南》（Why Zebras Don’t Get

Ulcers：A Guide to Stress, Stress-Relate Diseases, and Coping ）。

费米的“小发现”和复杂性理论的未来

S.斯特罗盖茨

Steven Strogatz

1942年12月2日，在芝加哥大学一个室内网球场进行的秘密实验中，费米（Enrico Fermi）建造了第一个自持的链式反应堆，迈出了原子弹发展的关键一步。[58] 费米也因此总被人想起，至少公众是这样记住他的。不过在科学家中间，他赢得尊敬却是因为他宽广的胸怀。费米也许是同时站在理论和实验物理学两个巅峰的最后一人。布罗诺夫斯基（Jacob Bronowski）[59] 曾有过挑逗性的评价：“在我印象中，他是我见过的人当中最聪明的——当然，也许是除了一个人以外的最聪明的人。”“他个头小，结实而有力，像个运动员，而且他心里总有明确的前进方向，仿佛能把事情看透。”

就在1954年去世前，费米从物理学家通常说的一个消遣问题得到了很大的乐趣。那是一个简单而优美的问题——没有讲什么太现实的东西，不过是探索他一直好奇的一个基本问题的方法。现在，他的机会来了。这时，费米在洛斯阿莫斯访问，正在找理由来试验新的MANIAC——世界上第一台超级计算机。那机器在他看来就是挡不住诱惑的一辆赛车。

费米同帕斯塔（John Pasta）和乌拉姆（Stanislaw Ulam）一起工作，他们想让机器模拟32个粒子组成的弹性链的数千个变量。整个体系代表一个理想化的通过化学键束缚的一维原子晶格。对于小振动，化学键表现出线性行为：如果原子分开的距离增大1倍，那么把它们拉回来的力也需要增大1倍。所有传统物理学都建立在这样的近似基础上。但费米知道，如果振动大了，真实的化学键行为应该是非线性的，没人知道那种情况下会发生什么事情。那时的数学给不出答案：没人能解那么多变量的非线性系统的方程。

当然，问题也在这里。费米构造这个问题是因为传统方法对它无能为力。现在，有了MANIAC的帮助，他和他的合作者就要照亮经典物理学的那个非线性的黑暗角落了。他们的发现令人震惊。他们原想干扰静态系统时，非线性最终会使系统热起来——就是说，使系统退化到充满随机的状态，所有可能的振动模式具有相等的强度。这是热力学告诉我们的事情。但是，计算机却说不。经过很长时间以后，粒子几乎又回到了起点。这个奇怪的结果使人类第一次发现，非线性可能是惊人秩序的一个来源。非线性带来混沌，非线性又把它夺走。[60]

费米为那个回声现象感到高兴——据乌拉姆回忆，他亲切地称它是一个“小发现”——不幸的是他没能活着看到发表的结果。失去了费米，帕斯塔和乌拉姆在写那篇文章时也没愉快过，他们默默地计算，写进洛斯阿莫斯的内部报告。10年以后，它才作为费米著作的一部分发表出来。

费米—帕斯塔—乌拉姆问题在20世纪50年代初一定显得特别孤单。那时候，物理学都是关于量子电动力学的，没有人考虑跟经典力学一样陈旧而枯燥的事情。那个问题不是已经彻底考察过300年了吗？费米却不同，他认识到这个题目还几乎没人碰过；所有非线性问题跟从前一样令人困惑。在今天看来，不论对这个问题的选择，还是为了解决它而设计的革命性的计算机实验，都说明费米是多么有远见。经过适当推广，他的远见实际上更符合我们的时代，也适合未来50年的非线性动力学。

1953年，非线性动力学几乎不能把握涉及两个耦合振子的问题，更不用说更多的振子了。振子是工程师的东西，在前半个20世纪里，工程师们设计了很多应用非线性特性的工具：为最早的收音机和电视机供电的真空管，雷达和通讯的锁相线圈，精密光学和眼科手术的激光。所有这些发明都依赖于自持非线性振子——特别是依赖于它们相互同步或与输入信号同步的倾向。可是这些技术所用的振子很少，一般只有一两个。大量振子的排列是不可能的，那时还没有预言大量振子的集合行为的数学。

能处理大量相互作用粒子的学科只有统计力学，这个物理学分支的产生原是为了解释亿万分子组成的气体的行为。费米是统计力学的大师，很清楚它在热力学平衡系统发挥的作用。遗憾的是，非平衡现象完全是另一码事。这也正是费米-帕斯塔-乌拉姆模拟所产生的令人惊奇的

结果：系统不像人们希望的那样固定在平衡状态。普通的统计力学在这里失去了意义。

在费米实验后的50年里，非线性动力学和统计力学都成熟起来了——还有一定程度的重合。过去几十年的几个伟大的理论成果，利用两个学科的技术取得了惊人的进步。物理学家费根鲍姆（Mitchell Feigenbaum）运用统计物理学的重正化群方法（赢得过诺贝尔奖的方法）证明，存在一定的普适定律，决定着从规则行为向混沌行为的转变。他的预言很快在诸如心房、化学反应和半导体等众多系统得到了证实。理论生物学家温弗里（Arthur Winfree）证明，巨大的生物学振子网络的同步化现象令人联想起相变，就像水在临界温度下突然凝结成冰。复杂系统的其他重要模型——考夫曼（Stuart Kauffman）的基因网络、巴克（Per Bak）的自组织沙堆、霍普菲尔德（John Hopfield）的人工神经网络——都通过统计力学与非线性动力学的结合得到了说明。

非线性动力学是跟着由几个组织原理所决定的一个简单逻辑结构进步的。最重要的一个原理是，小系统比大系统更容易。最早认识的一类非线性问题就是那些只包含两个变量的系统——如摇动的摆，其状态完全由位置和速度来刻画。如果这两个量确定了，我们就能预言它在未来任何时刻的准确位置。

三个变量的问题会野蛮得多：它们可以是混沌的。混沌意味着在确定法则统治下的系统也可能以随机的、似乎不可预言的方式运动。根据洛伦兹（Edward Lorentz）和其他混沌学者从1960年到1985年的研究，我们认识了这类奇异的行为方式的普遍性，解释了它们的一般特征。从生态系统中的种群涨落到水龙头无规则的滴水，混沌可能意外地出现在任何地方。混沌很快就在实践中应用开了，如编密码甚至作曲。

接着，非线性动力学的前沿转向了更大的系统，一般是大量相互作用单元构成的网络。在这方面，涉及大量耦合振子的费米-帕斯塔-乌拉姆问题更有现代意义了。在这类耦合系统中，还有进一步的组织原则将我们引向一些最容易处理的问题。有些原则是针对网络的组成单元而言的：例如，振荡单元的集合行为比混沌单元的集合行为更容易预言，同样的单元比多样的单元更容易预言。另一些原则是关于单元的联络方式的：规则或随机结构的网络比精巧联络的网络更容易。总结起来，这些启发指引我们去研究那些以离散晶格或其他简单排列方式耦合的全同振

子的系统。在过去几年里，非线性动力学就是这样发展的。几个热点问题是关于多个耦合单元系统的同步化，如激光系统、神经元系统和所谓约瑟夫森节的超导系统。除了离散晶格外，有些研究者在关注连续介质中的模式形成，如流体、化学反应、神经和心脏组织。也许最激动人心的还是关于螺旋波动的动力学，它可能关联着心室纤维颤动（最危险的一类心率不齐）。

跟这些巨大非线性系统相似的，是圣塔菲研究所的研究者们所谓的“复杂性适应系统”——一些想象的由千百万个竞争的有机体、化学物质、公司或证券交易所组成的系统，每个系统都适应各自的环境，从而也改变其他系统的环境。这些系统的计算机模拟模型显然是推测性的，然而它提供的线索却让我们隐约看到了自然选择对一大堆问题的影响：生态系统的恢复能力、生命的化学起源、市场中公司之间的竞争以及股市的涨落。

从许多方面说，这些达尔文式的模型都是费米-帕斯塔-乌拉姆研究的智力后代。我们发现它们对计算机的态度是多么相似：不把计算机作为玩弄数字的机器，而把它作为一种解释的工具，一盏照亮我们在黑暗中摸索的明灯；我们发现复杂的非线性系统也同样享有意外秩序的魅力；我们发现这些模型都缺乏现实性。正如费米的一维粒子链是晶格的刻意简化，今天计算机模拟的复杂性适应系统也是真实生态系统和真实市场的粗略漫画。现在看来，那是不错的策略，但在未来的几十年，这个领域还要成长。我们面临的挑战是，寻求能囊括更多实在而不失远见的方法。

我们将面临的第一道障碍是刻画复杂网络的联络。现有的模型具有理想化的随机或规则的拓扑，而我们需要知道真实网络是如何组织起来的。如果我们想知道大脑如何计算，细胞为什么癌变，那是基本的要求。在过去的三年，我们已经开始研究一些网络的具体结构，包括血液网络、神经系统、电力网络和互联网络。令人惊奇的是，这些网络尽管千差万别，却有着某些普遍的结构形式。

例如，它们都表现出小世界现象（大家熟悉的所谓6度分离）：几乎所有的节点对都通过一条很短的链相连接。[61] 而且，每个节点连接的数目倾向于幂律分布，那种分布曲线的尾巴比正态的钟形曲线要大得多。这意味着，绝大多数节点没有很好的联络，但是也有为数不少的巨

大的联络中心——像网上的雅虎或者生化反应系统中的ATP。这些拓扑特征一定会影响网络所能支持的集合动力学的类型——它们对随机破坏或恶意攻击的反抗，它们传播扩散或支持整体计算的能力，等等。但是我们现在还一点儿也不知道如何将网络的拓扑与它的整个动力学联系起来。

实际上，在自然科学以外，我们几乎还没有好的动力学模型。尽管网络遍及生物学、社会学和经济学，但是对决定基因、个人和公司之间相互作用的法则我们还知道得太少。这是我们在未来50年需要清除的第二道障碍。如果不能从数据推测局部的动力学，我们的复杂系统模型就永远不可能超越漫画式的描绘。

1952年，霍奇金（Alan Hodgkin）和赫胥黎（Andrew Huxley）根据他们对乌贼神经轴突的电生理学测量，重构了神经动力学，这是从数据推测动力学的一个经典例子。他们有很好的条件做控制实验，在实验中，神经隔膜上的电压可以调节为任何数值，然后维持在那个水平。通过测量作为电压函数的穿过隔膜的钠、钾和其他物质流，他们导出了单个神经元的非线性动力学的准确图像。

困难的是，这种策略当然并不总是充分的。在其他情形，它也许不可能在一系列希望的水平上把握任何变量。不过，还有更多的间接方法能做出这些必需的推测。例如，我们考虑控制细胞行为的遗传网络。现在我们有了DNA片段，能同时测量作为时间函数的几千个不同基因的行为，可我们还是不知道哪些基因在相互交流，也不知道它们以怎样的定量方式影响彼此的活动。所有信息都多少反映在DNA片段的数据上，可是我们不知道怎么破译那些密码。如果我们能发展根据多个时间序列的测量来推测动力学的系统方法，那进展一定有着巨大的意义——不仅对生物学，也对社会学和经济学。

刻画联络特征的问题比推测动力学的问题容易得多。不过，就算两个难题都解决了，我们还会迎面碰上一个最根本的障碍。一旦投身去探索千万个相互作用变量的非线性系统，我们就不可能躲避这个障碍。它在费米-帕斯塔-乌拉姆问题中就以原始的形式显露出来了，说实话，没人知道该对它做些什么。我们总是忽略它，或者随意地改变问题的提法，但有时却不得不直接面对它。

简单说，困难在于我们的大脑只能想象三维的东西。进化已经那样地将我们的大脑网络固定下来了。有了计算机辅助训练，我们也许能模糊地感觉更多的维，但我想大概还是画不出千万维的图像。

为什么这是个难题呢？因为，自100年前庞加勒（Henry Poincaré）在引力的三体问题中发现混沌以来，几何就成了我们在非线性动力学中的最佳伙伴。记住，非线性方程一般不能以封闭的形式求解，所以代数公式没有用了。但庞加勒证明，我们不需要公式。画出正确的图形，我们就能理解非线性系统的许多关键的定性特征。庞加勒的方法为每一个状态变量赋予一个坐标轴，如果只涉及两三个变量，我们可以很容易把动力学画出来。但是对今天千万个变量的问题，我们却手足无措了。几何方法仍然有效，利用更抽象的推理形式也取得了一些进步，但不把图像直接画出来，我们对动力学还是茫然的。

因为这一点，我们长久没能从理论上认识湍流。尽管控制方程已经知道100多年了，我们却看不出它们的解在庞加勒定义的“状态空间”里有什么表现。特别是，我们画不出吸引子——长期动力学的本质——因为状态空间是无穷维的。

你可能会说，我们早就把那个难题解决了。在气体或磁体的统计力学里，状态空间有阿伏伽德罗常数（一个23位数）个维度，我们不也很好认识了那些系统吗？是的，不过那只是在它们处于热力学平衡的时候。然后我们知道长期行为的统计性质，那是平衡分布告诉我们的；这也正是我们现在困惑的原因。我们不知道复杂系统的长期统计行为，因为它们是远离平衡的。又因为画不出系统的吸引子，我们不知道该做什么。

正如费曼（Richard Feynman）40年前讲的，“下一个人类智慧觉醒的伟大时代，很可能产生理解方程定性内容的方法。今天我们还做不到。今天我们还看不出水流方程包含着我们在两个旋转的同心圆柱之间看到的那种理发店店招式的涡旋。今天我们还看不出薛定谔方程是否包含着青蛙、作曲家和道德。我们不知道是否需要一个高高在上的上帝那样的角色。所以，不论是与否，我们都可以牢牢把握自己的意见。”[62]

假如我们有一天走进那个伟大的觉醒时代，我们需要从多维的魔影里解脱出来。计算机可能是我们的救星。一旦它们有了惊人的智能，应

该可以画出任意维的图形。在运行我们的模拟时它们就已经在做那种苦役了；也许有一天它们还能从复杂系统中抽象出自组织的规律。

上面的猜想引出一个更大的问题：如果哪天计算机比我们更擅长理论科学，我们还能从中得到乐趣吗？如果计算机以我们能理解的定性方式表达它们的观点，那它们就像残疾人的假肢，不过是自己躯体的延伸，在哲学上并不比电子显微镜产生更大的影响。但是，假如它们使我们茫然无知，那就像过去的大圣人，神秘莫测，还常常产生混乱。在有的数学中，那样的事情已经发生了。有些定理就是计算机证明的，但是因为证明包含了实在太多的相互交织的小问题，没人能够检验。在深蓝与卡斯帕罗夫（Garry Kasparov）的对局中，有些棋的下法就有着同样的特点。

我想知道，复杂系统研究的未来是否就是那样。我们也可能像旁观者那样停下来，跟不上我们制造的机器，只能为它那些惊人的结果感到目瞪口呆。

费米也许是第一个产生这种可怕感觉的人。他发明的计算机实验为科学研究开辟了全新的道路。那样的事情之所以可能，是因为同时代的冯·诺伊曼（John von Neumann，最早的高速计算机建造者）的工作，而那似乎是理所当然的——正如布罗诺夫斯基讲的，冯·诺伊曼“是我认识的最聪明的人，没有例外”。

斯特罗盖茨（**Steven Strogatz**）

斯特罗盖茨（Steven Strogatz）是康乃尔大学应用数学中心教授。他写过一本很受欢迎的教科书《非线性动力学与混沌：在物理学、生物学、化学和工程中的应用》（Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering），即将出版的还有一本普及读物《同步》（Sync）。他在人类睡眠和生命节律、涡波、耦合振子、同步萤火虫[63]、约瑟夫森节以及小世界网络等方面的开创性研究发表在众多出版物和广播等媒体，包括《自然》《科学》《科学美国人》《纽约时报》《纽约人》BBC（英国广播公司）电台和美国CBS（哥伦比亚广播公司）新闻。

生命是什么

S.考夫曼

Stuart Kauffman

在这个分子生物学凯歌高奏的时代，在第一个人类基因组草图产生的时候，人们可能认为我们应该知道怎么回答这个问题了：生命是什么？但我们不知道。我们零星地知道一些分子机器、新陈代谢循环、遗传网络路线以及膜生物合成的方法，但我们不知道一个自在活泼的细胞是怎么活起来的。问题的核心还是一个谜。

我想，我们将在未来50年回答那个问题，而答案需要物理学和化学的显著改变，更不必说生物学了。实际上，就生物学而言，对生命是什么的基本问题的认识，将开辟一个广义生物学的时期，而不仅限于我们唯一知道的关于地球生物的生物学。我们将有能力问，是否存在统治宇宙任何地方的生物圈的定律？

并不是说聪明的头脑拿不出一个答案。最有名的也许是跟本文标题相同的一本书，发表在战火纷飞的1944年，作者正是大物理学家薛定谔。[64] 不过，薛定谔的精彩著作却把另一个问题作为核心，其答案似乎并没有回答他的（和我的）题目呈现的问题。他那个核心问题是关于生命系统中惊人的秩序的起源。薛定谔在回答中指出，秩序不可能来自统计平均，我们都知道那里的涨落尺度大概是粒子数的平方根。他利用新近的X射线诱导的基因突变率，正确认识到一个基因最多可以由数百到数千个原子组成。均方根尺度的涨落不能表现生命的那些可遗传特征。于是，薛定谔做了大胆的思想飞跃：他指出，生命的秩序需要稳定的化学键——特别是共价键——它们依赖于量子物理学，而不是经典物理学。接着他指出，简单的晶体“说”不出多少东西，因为单位晶体的结构知道了，它的全同重复也不能增添“说”得更多的力量。所以，薛定谔精明地把希望寄托在非周期的晶体，其具体结构将包含确定生命发展的微码。他做对了——9年以后，沃森和克里克才发现非周期DNA固体的结构；大约再过10年，我们才认识到这种微码起着遗传密码的作用。

不过，如果说薛定谔敏锐地预见了生命秩序的起源，那么，他是否

也回答了他的问题，生命是什么？我想没有。我不能立刻说我为什么那么相信这一点。这篇文章是为了解释关于这个问题的另一种尝试。

我从一个不同的图像和问题说起。想象一个逆着葡萄糖梯度游动的细菌。不必为细菌赋予什么意识，我们都会毫不犹豫地说，它在寻找食物。就是说，细菌在环境中为自己而活动。我把能在环境中为自己活动的系统叫做自治行为者。一切自由活动的细胞和有机体都是自治行为者。但是，细菌“只不过”是分子以一定方式组织起的一个物理系统。所以，我的问题变了，不是“生物学秩序的起源是什么”，而是“什么物理系统才能成为自治行为者”。

我先把我暂时的答案告诉你们：我相信一个自治行为者一定是能自我繁殖并能实现至少一个热力学功循环的物理系统。就细菌而言，它旋转的鞭毛“发动机”因为反抗它周围的液体介质而做功。细菌能自我繁殖，而且能在游动中做功。

这就立刻引出一大堆需要注意的问题。当我描述细菌在为自己活动时，其实是在用我们描写人类活动的语言。关于人类，“做”“活动”、“目的”、“选择”等概念都是我们熟悉的，深深嵌在维特根斯坦（Wittgenstein）的语言游戏 [65] 或我们所有人的生活方式中。即使我们要怀疑是否能把“为自己活动”的概念恰当地用于细菌，同样会产生一个令人困惑的问题：活动、做和目的等用于生活在物理世界中的人的字眼是从哪里来的？我们自己也不过是物理系统。所以，我要跳过关于其他思想的哲学问题，以及我们想把语言游戏向生物世界延伸多远的问题。我个人的意见是，我们确实把我们的语言延伸到了自由活动的生命，甚至单细胞细菌，更不用说一对筑巢的小鸟和追逐你扔出的小木棍的狗。这也证明了我关于细菌的问题。如果我是对的，那么，自治行为者就是我们关于“做”和“行为”，从而也是我们“为自己活动”等语言游戏的充分而必要的条件。

在我的自治行为者的试验定义中，我也许无意间发现了生命的适当定义。行为者与生命当然有着相同的外延，尽管难免存在边缘的例子——如杂交物种等——它们显然也是生命，但不能繁衍后代。我不会顽固地坚持自己的自治行为者的定义是充分的，也不认为它一定能满足生命——但我想它可能是的。

现在简单说几句科学中的定义循环。想想牛顿著名的f=ma，问问自己，如何独立于m来定义f？力是使质量加速的东西，惯性质量是反抗力的加速的东西。庞加勒主张（我赞同这个观点，但并不是所有物理学家都同意），f=ma是一个定义循环，力和质量是彼此相互定义的。定义循环没有妨碍天体力学的巨大成功。听达尔文的宣言：自然选择选择适应性，而适应性是能使生命更容易产生后代的特征。这也是一个定义循环，它也没有阻碍进化论光大达尔文的思想。

我的自治行为者的定义也是一种定义循环，同时还跳进了新的语言游戏——关于做与行为的游戏。“自我繁殖”和“功循环”都是独立于自治行为者而定义的，就这点说，我的定义并不真的是定义循环。但是，进一步把自治行为者等同于为自己行为的能力，却是定义循环。当然，那并不意味着定义没有优点和科学意义。我曾力图去认识和解析包含在我那简单的自治行为者定义里的东西，结果走进了一条概念分析的小路，说明它至少包含了一些令人感兴趣的东西。别人一定会问它最终是不是有用。

我们粗略地来看一个可能的化学系统，它既能自我复制，也能实现功的循环。首先，系统处在某个能自我复制的部分。它可以是6个核苷酸的一段DNA，能激发两段3个核苷酸的序列相结合，它们一旦结合起来，就转变为跟原来一样的6个核苷酸的DNA；它也可以是31个氨基酸的序列，能激发两个片段的结合——15个氨基酸的一段与17个氨基酸的一段——形成跟原来的一样的31个氨基酸的序列。顺便说一句，斯克里普斯的加德里（Ghadiri）小组成功实现了蛋白质的自我复制，决定性地证明了分子的复制不必建立在DNA和RNA类分子的模板复制的基础上。

至于自治行为者定义中的功循环，可以通过某些假想的化学过程来实现。在那些反应中，焦磷酸盐PP分解为两个一磷酸盐P+P，同时损失一定的自由能。那个自由能耦合在额外复制的6个核苷酸的DNA或31个氨基酸的序列中。这里，“额外”的意思是，假如复制反应没有结合焦磷酸盐产生的自由能源，就不会复制出那么多的分子。一旦焦磷酸盐分解了，我想象可能还有一个自由能的来源——那些吸收光子并受光子激发的电子。当电子回到它原来的未激发状态，这个自由能被用来驱动焦磷酸盐的再合成，超过没有电子能量激发时所能达到的平衡浓度水平。功循环的存在基于这样一个事实：在从PP到P+P然后又回到PP的反应循环

中，磷酸盐使反应循环“转了一圈”，而不是达到平衡。这是实现热力学功循环的一个分子发动机。

我那个试验性的定义有许多显著的特点。首先，系统只有在远离化学平衡时才成立；行为者的活动是非平衡概念。其次，系统是一类新的完全实在的非平衡化学反应网络，结合自我复制与功循环的网络。只是我们过去从来没有把二者结合起来。不过我们现在可以用实验来考察它们了。第三，物理学家安德森（Philip Anderson）曾向我指出，总系统通过额外合成6个核苷酸的DNA或31个氨基酸的序列，储存了供后来修正错误的能量，我们细胞里的DNA修复酶就做着这样的事情。

写到这点为止，本文还算是正统的科学。下面的事情会越来越奇怪，也可能有更多的问题。

第一个问题是更仔细来分析“功”的概念。对物理学家来说，功不过是力作用一段距离。用球棍击球加速的时候要做功，构成加速运动的正是这些小小的ma项的总和。

但是功真就那么简单吗？在任何具体的功的情形，力是“有组织”地作用在物体上做功的。物理学家在他们的初始和边界条件中隐藏了过程的组织思想。可是初始和边界条件又从何而来呢？这是物理学家通常没有回答的问题。就分析炮弹射出大炮的运动说，完全可以忽略以前的事情，只考虑炮、高程、火药、炮弹的重量、气流等初始条件，然后进行下一步计算。但是，造炮、装火药、发射等都需要做功。这样一路追溯下去，完成一颗炮弹的发送大概要回到宇宙大爆炸。生物圈讲的是过去35亿年来不断出新的初始和边界条件的产生，物理学把“现在”这个系统孤立出来，不足以回答为什么会进化地涌现那样常新的初始和边界条件。

为了认识功的概念有不完备的地方，第二个办法是要认识到一个孤立的热力学系统——例如在热力学上孤立的一个盒子里的气体——是不能做功的。但是，假如用薄膜将盒子分为两部分，则一个部分可以对另一个部分做功。例如，第一部分的压力较高，薄膜就会向第二个部分移动，从而对它做力学的功。这样的功不可能在宇宙实现，除非宇宙也至少分为两个区域。另外，宇宙的薄膜又从哪儿来呢？

我感觉最合我意的“功”的定义出自阿特金（Peter　Atkin）关于热力

学第二定律的一本书。他在书中指出，功是一样“东西”——即能量在约束条件下的释放。考虑一个内有活塞的圆柱，活塞与圆柱顶之间充满压缩的工作气体。气体可以膨胀，对活塞做功，将它压下。约束是什么呢？显然，是圆柱、活塞和活塞在圆柱中的位置，以及约束在其中的气体。那些约束从哪儿来呢？我们来看，做圆柱要做功，做活塞要做功，然后，把气体和活塞放进圆柱也要做功，于是，我们遇到一个以前没有留意过的有趣的新循环。看来，为了产生约束我们需要做功，而为了做功我们需要约束！

实际上，当我们试图解开薛定谔包裹在他那非周期性晶体微码中的信息概念时，为了“说”出有物理意义的东西，只有通过对能量释放——能量在释放时形成功——的约束进行重新组合。更重要的是，释放出来做功的能量，可以对能量的释放构成更多的约束，那些约束又做更多的功，功又形成更多的约束。注意，这些都不是我们过去在物理学和化学中学过的概念。人们现在才开始逐渐感到所有这些加在能量释放的约束——作为功，它们能对能量释放构成更多的约束——密切联系着某个充分的关于过程组织的理论。那样一个理论，我们目前甚至连轮廓也没看见，当然，那轮廓也不会包含在信息论里。

我说这一点也不仅是文章写作的需要。分裂的细胞就做着我上面说的那些事情。例如，细胞做热力学功，从脂肪酸和其他组成要素形成脂类分子。然后这些脂类分子可以回落到低能量的结构——一个双脂层，能形成被称作脂质体的空腔“气泡”。实际上，细胞膜就是这样一个形成空心球的双脂层。这时候，膜本身被细胞拿来改变约束。考虑一对假想的有机小分子A和B，它们可以经历三种完全不同的化学反应：

1）A和B能反应产生C和D；

2）A和B能反应产生E；

3）A和B能反应产生F和G。

每个反应都有反应坐标[66]。这样，当A和B反应产生C和D时，化学势阱和势垒将引导着分解的分子A和B形成产物分子C和D。同样，势阱和势垒引导着反应形成E、F和G分子。这样，化学势阱和势垒就构成A和B反应的约束。进一步，假定A和B从细胞内部的水溶液分离出来转

化为细胞膜。一旦进入细胞膜，A和B分子的运动模式，如平动、转动和振动，都将发生改变。这反过来又会改变将A和B与C和D（或E、或F和G）分隔的化学势阱或势垒的深度。这样，细胞膜的形成（做了热力学功的）改变了A和B所能产生的化学反应的约束。而且，细胞也做热力学功，将氨基酸结合成为蛋白酶，那正是约束A和B并将其转化为C和D（而不是E、或F和G）的物质。这样，细胞做的热力学功形成了特殊路径的能量（这里是化学能）释放的约束。能量一旦释放出来，又做功形成新的约束。

我说的是正确的，但没有以通常的方式来谈。实际上，细胞所做的是我后来所谓的“繁衍功”——它将一个地方的有组织的能量释放，与其他地方的约束的形成和有组织的能量释放联结起来，形成一个闭合的过程，细胞就通过这样的过程完成大概的自我复制。任何正在分解的细胞都在实现过程的这种组织，一个细胞分裂为两个、四个、以至千千万万，最后形成一个生物圈。不过令人惊奇的是，我们还没有什么语言——至少，据我所知还没有那样的数学语言——能描述这个闭合的过程。过程的这种自我繁衍的组织包含在自治行为者的定义中。注意，尽管我讲的这些没有违背物理学定律，但物理学和化学似乎也不存在讨论它们的语言。我们还应该看到，“自治行为者”的试验性定义开始显现出多么丰富的内容。不过，我还是凭直觉感到，细胞似乎多少表现出一种不为我们的信息概念所把握的组织形式——那个概念从来没有考虑过在现实的物理世界真实发生的任何事情上构造什么约束。实际上，生物圈的细胞增生繁殖的协同演化过程，也是更加微妙和精致的演化路径不断出现的过程——在约束下通过那些路径获取和释放自由能；而约束的路径又产生新的约束和结构，繁衍的功从约束中汲取新的功，形成更加微妙的结构。我们只需要想想热带雨林。雨林的繁茂没有任何中心力量的作用，而它却产生了蓬勃的相互交错的多样的连环过程。是的，生物圈产生了有组织的功的繁衍的连环模式。但是，如何把这个概念数学化，我连怎么起步都不知道。

我不知道如何把这个概念数学化的原因，也许说明了本研究最奇异、最令人沮丧、也可能最深刻的一面：生物圈可能实际正做着谁也不可能预见的什么事情。如果真是那样，牛顿、爱因斯坦、玻尔和玻尔兹曼为我们指引的科学道路就有局限。生物圈可以持续改变它的“相空间”。我不知道有没有能描写这种过程的数学框架。我还想指出，如果我说的是正确的，它将给概率的频率解释带来巨大的灾难——在概率论

中，我们实际上可以事先谈论所有的可能性。

这个问题联系着达尔文所说的预适应性。不过我们先说说一般的适应性。我们心脏的功能是什么？当然，我们心脏的功能是输运血液。它也产生心跳，但那不是它的功能。所以，心脏的功能只是它所产生的结果的一部分。达尔文会说，自然为我们心脏选择的功能是输运血液，不是产生心跳。就是说，为了了解上面讲的A和B或生命的任何部分的功能，我们必须在整个环境中去认识整个生命。自治行为者是从绝对的整体论来认识它们的。

现在谈预适应性。达尔文的观点是，生命的某个部分产生的结果，在正常环境下可能没有选择意义，但在特殊环境下也许有意义从而被选择。注意，几乎所有的主要适应性和所有（或多数）一般适应性都是这样产生的。于是才出现了飞行、视力、听力、肺、咽喉，等等你能想到的一切。

我的困惑是这样的。你认为你能预先罗列所有可能的（例如，针对现存的生命形式）达尔文的预适应性吗？更正式地说，你能事先限定所有可能的预适应性吗？我还没有遇到谁认为那答案是肯定的。我们甚至连这个预适应性的清单如何开头也不清楚。现在，我深信这种说法是对的，但不知道它是经验总结还是数学推论，也不知道现在如何去证明。

现在我们来听听小松鼠格特鲁德的故事。她生长在2350万年前，经历过一个孟德尔的优势基因突变，于是留下两层皮肤，每层都从腕延展到踝。格特鲁德那么丑，没有同族的伙伴同她一起说话，一起吃东西，一起玩儿。一天，格特鲁德正在一棵玉兰树上沉思，被邻近松树上的大枭贝莎看见了。呀，多美的午餐！贝莎想着，张着利爪闪电般俯冲下去。格特鲁德吓坏了，“嘎嘎”呼叫着张开双臂跳下树去，格特鲁德飞起来了！贝莎疑惑地眼睁睁望着她逃走了。于是，格特鲁德成了她的种族的英雄，1个月后跟一个漂亮的小伙儿在举族同欢的盛大典礼上结婚了。因为那个起主导作用的基因突变，不久诞生了许许多多的长着两层皮肤的松鼠宝宝，我们今天看到的会飞的松鼠，或多或少是这样来的。

那么，你能事先告诉我们那褶皱的皮肤能在那天成为翅膀吗？也许。你能说细菌中某个模糊的分子突变能让细菌探出纤毛虫的钙质流体然后向它发起进攻吗？我想不能。更一般地说，我们正好没有先验的概

念能说所有可能的达尔文的预适应性是什么，也不能说所有可能的环境是什么。

现在请注意，这些不可预言的事情一刻也没有减缓进化的脚步。达尔文的预适应性时刻在产生。于是，生物圈似乎在做着我们不能预先说明的事情——不是因为量子不确定性或混沌动力学行为，而是因为我们事先一点儿概念也没有。

这反过来又意味着生物圈的相关可能性的空间——它的相空间——也是不可能预言的。于是，生物圈具有我们不可预言的创造力。这就跟牛顿给我们的谆谆教诲形成了尖锐的矛盾：在物理学中，一般可以预言所有可能性的集合——也就是相空间——然后借助定律与初始和边界条件来计算粒子在相空间里的前进轨道。

我想我们不可能在生物圈里事先描述这种相空间（所有可能性的空间）。物理学家可能会说，“不过，假如你以经典的方式来处理这个系统，总会存在关于系统（以某种方式孤立的）粒子的所有位置和速度的经典的n维相空间。”也许是那样的，可是，我们还不知道哪些相关的集合变量（如格特鲁德的翅膀）才是揭开生物圈秘密的决定性变量。于是，我们似乎面临着以前从未认识到的知识的极限。进化的生物圈做着不可预言的事情，我们还没有起码的概念。我想，技术的进步也是这样的：100年前没人能预见互联网。

有趣的是，我们不能预言未来技术的事实，假如正确的话，粉碎了统治着当代经济学理论的核心：“竞争性一般均衡”——它一开始就假定我们能预言所有可能的货物和服务，然后证明市场干净了——即所有货物都按合同价卖给买家了。但是我们不能事先知道所有可能的货物和服务，所以那个统治的经济理论一开始就错了。它把焦点放在清理市场，忽略了一个在全球经济中起着决定作用的事实：货物和服务的多样性已经跟物种的多样性一样暴涨了。为什么？

物理学的基本定律是时间可逆的。我们已经相信，时间的出现是因为热力学第二定律的不可逆性。我不反对这一点，不过我好奇的是生物圈中源于达尔文预适应性的隐藏在过去与未来的区别中的时间箭头。那似乎不是第二定律在起作用，而是达尔文的变异与选择。

还有最后一点。宇宙是不会重复的，也不是各态历经的。考虑200

个氨基酸的蛋白质数目。因为有20种氨基酸，那种蛋白质的所有可能的数目就是20的200次方，即大约10的260次方。现在，已知宇宙间的粒子数大概是10^{80}，我们可以问问自己，从宇宙130亿年前大爆炸以来是否造出了所有可能的蛋白质。快速的化学反应需要1飞秒（10^{-15}秒），于是，自大爆炸以来在1飞秒的时间尺度上可能发生碰撞的粒子对的数目为$10^{80} \times 10^{80} \times 10^{33}$，等于$10^{193}$。这是一个大数，但同200个氨基酸的蛋白质的数量$10^{260}$相比，却是微不足道的。于是，即使宇宙的一生除了造蛋白质外什么都不做，也需要重复10^{67}次才可能造出所有可能的200个氨基酸的蛋白质。因此，一旦超越原子核的水平，在有机分子的复杂性水平上（更不必说生命、法律系统或雪弗莱汽车），宇宙在很大程度上是不重复的。这意味着当格特鲁德实现她的第一次飞行时，她改变了宇宙的物理和分子的演化。许多或大多数自治行为者的行为也是如此。我们是走在独特轨道上的过客，我们能产生实在的影响，不是吗？美国人赶在苏联的前面登上月球，在上面留下了一点儿东西，从而改变了太阳系的轨道动力学。

我们从起点出发已经走了很远——开始的时候，我们想知道什么样的系统能为自己活动。我不能确信自己是否在无意间发现了充分的生命的定义，不过我还是倾向于认为我可能发现了。而且，我相信未来50年能在实验中看到产生更多这样的系统。当它们彼此协同演化时，我们将惊奇地看着眼前发生的事情却不能预言将要发生什么。正如《侏罗纪公园》里的“混沌家”的著名台词说的，“生命发现了一条路。”他没有补充一句，我们一般不可能预先有任何线索说明那条路将是什么。生命天生是开放的，它的认识需要将物理学和化学提高到一个新的水平，在那里，从预言的范畴说，未来是开放的，而不是可以预料的。

考夫曼（**Stuart Kauffman**）

考夫曼（Stuart Kauffman）是宾夕法尼亚大学生物化学荣誉退休教授，理论生物学家，研究生命起源和分子组织的起源。他是麦克阿瑟荣誉学者和圣塔菲研究所终身教授。25年前，他创立了表现出一类自组织（他称为“免费的秩序”）的随机网络的考夫曼模型。考夫曼博士是Bios集团公司的主要创建伙伴和科学总负责人——公司将复杂性科学应用于商务管理问题。他是《秩序的起源》（Origins of Order ）和《探索》（Investigations ）的作者，还写了《宇宙为家》（At Home in the Universe: The Search

for the Laws of Self-Organization ）[67]。

第二部分　未来的实践

小摩尔定律

R.道金斯

Richard Dawkins

有过伟大贡献的人已经走得很远了，但有时为了自己高兴他们会走过了头。在评论《双螺旋》时，梅达瓦（Peter Medawar）[68] 写道，“假如谁愚钝到不能认识‘分子遗传学’的系列发现是20世纪最伟大的科学贡献，就根本不值得跟他来讨论。”他当然明白自己在做什么；他和他所评论的那本书的作者一样，在很大程度上是有资格高傲一点的。不过你也不必那么“愚钝”地去反驳他的意见。那么，英美人的那个以“新达尔文现代综合”出名的系列发现呢？物理学家为相对论或量子力学、宇宙学家为膨胀的宇宙，都能找到很好的骄傲的理由。“最伟大的”东西到头来是谁也说不准的，但分子遗传学革命无疑是20世纪——也是人类有史以来——最伟大的科学贡献之一。在未来50年里，我们能把它引向何方？或者，它将把我们带向哪里？到了那时，历史也许会证明，梅达瓦比他的同辈甚至他自己所想象的，离真理更近。

如果要我用一个词来总结分子遗传学，我会选择“数字的”（digital）。当然，孟德尔的遗传学，对基因在族谱中独立分配来说，是颗粒的，也是数字的。但基因的本性还是未知的，它们仍然是物质的东西，不断改变着数量、力量和品性，不可分割地跟它们外在表现卷在一起。沃森和克里克的遗传学是地道的数字的，连它的骨干即双螺旋本身也是数字的。一个基因组的大小可以用几十亿个基来度量，就像拿几十亿个字节来度量硬盘的大小一样精确。实际上，基与字节这两个单位可以通过一个常数因子相互转换。今天的遗传学纯乎是一门信息技术。正因为这一点，北极鱼的抗冻基因能复制到西红柿的身上。

在沃森和克里克联合发表他们著名的论文以来的50年里[69]，引发的知识爆炸跟任何一个受欢迎的爆炸一样，在以指数形式增长着。我想我说的就是爆炸。还有一个类似的更有名的爆炸的例子可以支持我的说法——例子来自传统意义的信息技术。摩尔定律指出，计算机的能力每18个月增强1倍。这是一个经验定律，没有公认的理论基础，尽管纳森·迈尔沃德（Nathan Myhrvold）提出了一个能巧妙地“自我证明”的候选者：“纳森定律”说，软件增长比摩尔定律还快，所以我们才有摩尔定律。不管基本的理由是什么，也不管它是多少理由的复合，摩尔定律在近50年的时间里都是成立的。许多分析家预计它还能继续成立50年，在人类事物中产生令人惊奇的影响——但那不是本文关心的事情。

我关心的是，在DNA信息技术里有跟摩尔定律相当的东西吗？最好的度量当然是经济的，因为钱是工时和机器消耗的最佳复合指标。几十年过去后，一定基本数量的钱能排序多少个基本单位的DNA呢？它以指数方式增长吗？如果那样，它多长时间增加1倍呢？注意，在这样的方式（这是DNA作为信息技术分支的另一个方面），DNA来自哪种动物和植物是没有区别的。在任何一个10年，排序技术和花费完全是一回事情。实际上，除非读了信息本身的内容，否则不可能知道那个DNA来自人，来自蘑菇还是来自细菌。

虽然选择了经济指标作基准，我还是不知道如何度量实际的花费。幸运的是我想到了问我的同事，牛津大学遗传学教授霍吉金（Jonathan Hodgkin）。我高兴地发现，他最近在为母校准备一个演讲时也做过同样的事情，他慷慨地把他对花费的估计寄给了我。他计算的是排列每一个碱基对（也就是每个DNA码的“字母”）所需要的英镑数。1965年，为细菌的5S核糖体的RNA排序时（不是DNA，不过RNA的花费是一样的），每个字母大约需要1000英镑。1975年，为病毒X174的DNA排序时，每个字母大约需要10英镑。霍吉金没能找到1985年的好例子。不过，在1995年，每个字母只需要1英镑就能排序*Caenorhabditis elegans*的DNA。那是分子生物学家非常钟爱的一种小线虫，不说名字也都知道是“它”。[70] 2000年人类基因组计划达到顶峰时，每个字母的排序花费大约是0.1英镑。为了说明真实的增长趋势，我把那些数字转换成“钱换来的东西”——就是说，一定数量的钱可以排序的DNA的数量。扣除通货膨胀的因素，我选择1000英镑为基数。我把每1000英镑排序的DNA碱基数（以千为单位）画在对数坐标上，这样做很方便，因为指数形式的增长将表现为一条直线（见后面的图）。

我必须强调，正如霍吉金教授做的那样，这4个数据点是在信封的背面计算的。不过它们确实令人信服地接近一条直线，说明我们DNA排序能力的增长是指数形式的。加倍的时间（或花费减半的时间）是27个月，大概可以跟摩尔定律的18个月相当。DNA的排序依赖于计算机的能力，在这个意义上（很大程度上），我们发现的这个定律也许应该归功于摩尔定律本身，这也为我那个玩笑的题目“小摩尔定律”找到了根据。[71]

我们永远也不能指望技术以如此的指数方式进步。我没有把图画出来，但是，假如看到飞船的速度、汽车的燃料价格或者摩天大楼的高度像指数那样增长，我一定会感觉惊讶。我猜它们以某种接近算术加法的方式增加，而不会在一定的时间内加倍再加倍。实际上，已故的伊万斯（Christopher Evans）早在1979年摩尔定律几乎刚开始的时候，就写道：

> 今天的汽车在许多方面都跟战后那些年的汽车不一样……不过我们可以设想一下，假如汽车工业在这些年里也像计算机那样发展，那么最新款式汽车的价格会便宜多少，效率会提高多少呢？……今天你大概可以用1.35英镑买一辆劳斯莱斯，它1加仑汽油能跑300万英里，它产生的马力还能驱动伊丽莎白女王二世号[72]。如果你对小型号感兴趣，还可以拿七八辆来放在针尖上。[73]

在我看来，空间探测很可能也像汽车那样不紧不慢地以加法的方式发展。我想起克拉克（Arthur C.Clarke）讲过的一个迷人的猜想，他的预言本领是令人难忘的。想象未来的一只飞船向着遥远的恒星飞去，即使以当时技术所能允许的最高速度飞行，它也需要好多个世纪才能到达那遥远的终点。在它还没走完一半的时候，一艘更快的飞船追了上来，那是100年后的技术产儿。于是，我们大概可以说，原来那个飞船本就不该忙着发出去。同样的道理，第二艘飞船也不应该发射，因为船上的人们注定会看着他们的重重孙子们坐着第三艘飞船从身旁飞快地掠过……其实，只要我们指出，没有前辈的慢的技术的研发，就不会有后来的快的飞船，悖论就解决了。有人认为，完整的人类基因组计划可以现在从头开始，而且只需要实际计划的部分时间就能完成，这样，原先的计划本该适当推迟一些时候。谁提出这样的建议，我也拿上面的话来回答他。

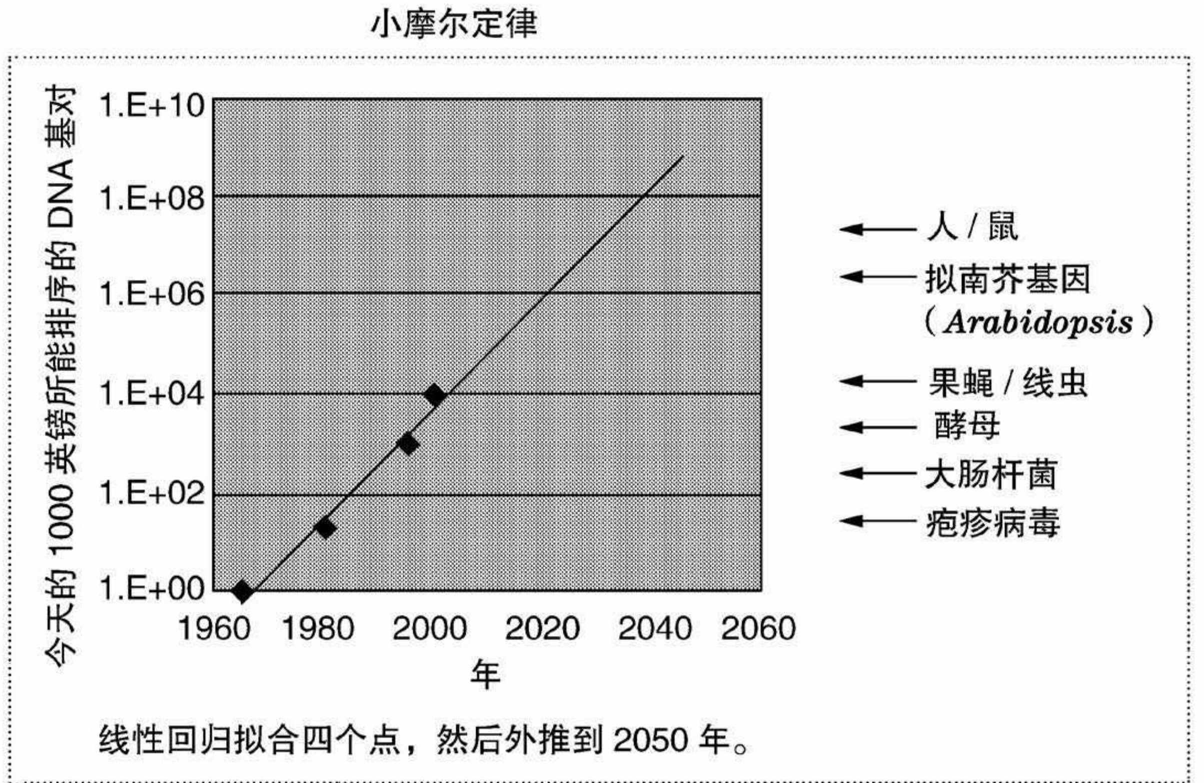
如果说我们的4个数据点勉强可以作为一个大概的估计，那么把直线外推到2050年就更不可靠了。不过跟摩尔定律类比，特别是，如果“儿子定律”确实与它的“父亲”有点儿关系，那么直线可能代表着某个合理的预言。至少，我们可以看看它能把我们引向何方。它告诉我们，在2050年的时候，我们可以用今天的100英镑（大约160美元）完成一个人的基因组排序。除了“那个”人类基因组计划，每个人都能负担他个人的基因组计划。种群遗传学家们将获得人类多样性的最终数据。我们有可能画出一个把世界上的任何一个人跟任何其他人联系起来的谱系图。这是历史学家苦苦追寻的一个梦。他们将利用基因的地理分布来重现几百年来的大移民和大入侵，追溯海盗的长船航行；他们根据基因去跟踪从阿拉斯加到火地岛的美洲部落，跟踪撒克逊人穿过不列颠，寻找犹太人散居的证据，甚至发现尚武好斗的成吉思汗们的现代子孙。

今天，胸腔X射线可以告诉你是否有肺癌或者肺结核。到2050年，用胸腔X光的钱，你可以知道每个基因的全部信息。医生给你的，不是大家一样的病历本，而是针对你个人的基因组的处方。那当然是好事情，但你个人的清单也惊人准确地预言了你的自然终点。我们需要知道那样的东西吗？即使自己想知道，我们会把个人的DNA图交给保险计算员、血缘律师或者政府吗？即使在仁慈的民主社会，谁也不会为那种前景感觉幸福。我们还担心，会不会有某个未来的希特勒来滥用这些知识。

这些担心尽管很重要，但也不是我这篇文章要说的。我要退回自己的象牙塔，关心更专业的东西。假如100英镑能排序一个人的基因组，那么它也能用来买其他任何动物的基因组。大小都差不多为几十亿个基的数量级，所有脊椎动物都一样。即使假定“小摩尔定律”在2050年之前就失效了（很多人都是那么想的），我们仍然可以放心地预言，我们每年在经济上能够排序的基因组，可以来自几百种典型的脊椎动物、几千种昆虫或其他非脊椎动物、几十万种细菌、几百万种病毒，或者数量令人不安地变化着的爬行动物或者开花植物。获得那么一堆杂乱的信息是一回事，用它们来做什么则是另一回事。我们应该如何消化它们？筛选它们？比较它们？使用它们？

一个不那么野心的目标是通过那些信息来最终把握系统发育的谱系。毕竟我们已经有了一个真正的生命谱系——实际发生的那个独一无二的生物进化分支模式。它确实存在着，也可以从理论去认识。我们现

在还知道得不完全。到2050年的时候，我们应该知道了——假如还不能知道，也只是数不清谱系大树上的细枝嫩丫，因为我们不知道物种到底有多少[正如我的同事迈（Robert May）说的，关于物种的数量，我们知道的至少也要差一到二个数量级]。



我的研究助理提出，2050年的博物学家和生态学家将随身携带一个小小的野外分类工具包，这样就不需要把标本寄给博物馆的专家去鉴别了。在便携式电脑上接一个精细的探头，把探头插入一棵树、一只新捕获的田鼠或者蝗虫，电脑会在几分钟的时间内检查几个关键的DNA片段，然后确定种属的名称和其他可能储存在数据库里的细节。

DNA分类已经发现了一些令人惊奇的东西。根据我过去做动物学家的经历，要我相信河马与鲸鱼的关系比与猪的关系更近，几乎是难以忍受的。这一点还有争议。到2050年，这个问题跟其他无数争论一样，总会有办法解决的。说它能解决，是因为那时候河马基因组计划、猪基因组计划和鲸鱼（假如到那时它们还没有被我们的日本朋友吃干净的话）基因组计划都要完成了。实际上，为了最后清除分类的不确定性，并不

需要把所有基因组都全部排列出来。

一个附带的好处，可能在美国产生巨大的影响：如果彻底弄清了生命的谱系，要怀疑进化的事实就更加艰难了。在这个问题上，化石相对说来没有多少意义。因为在现存的那么多我们能排序的物种里，几百种独立的基因相互证明，确立了那个真实的生命谱系。

有句话老被人提起，几乎说滥了，不过我还是要再说一次：认识一个动物的基因组并不等于认识了那个动物。根据布雷纳（Sydney Brenner，唯独这个人，我听说人们一直很奇怪为什么诺贝尔奖到现在还没给他 [74] ）的观点，从基因组“计算”动物，我想应该分三步走，一步比一步难。第一步根据基因的核苷酸序列计算蛋白质的氨基酸序列，这很难，不过现在已经完全解决了。第二步根据氨基酸的一维序列计算蛋白质的三维卷曲模式。物理学家相信这在原则上是可以做到的，但做起来很困难，而且更灵活的办法通常是制造蛋白质，然后看它会发生什么事情。第三步根据基因和它们与环境（多数由其他基因组成）的相互作用计算发育的胚胎。这是目前为止最困难的一步，但是胚胎学（特别是关于同源和相似基因的）进步很快，到2050年时有可能解决它。换句话说，我猜想2050年的胚胎学家可以把未知动物的基因组输入计算机，然后计算机模拟胚胎的发育，最后生出一个完整的动物。这本身算不得特别有用的成果，因为子宫或卵总是比电子计算机更廉价的计算机。不过它从一个方面标志着我们的认识已经完整了。而具体实现那样的技术也是有用的。例如，侦探可以根据他发现的一丝血迹用计算机生成一张嫌疑犯的脸——甚至从婴儿到暮年的一系列脸，因为基因是不会老的！

我还想，到2050年的时候，我梦想的那本“死者基因大全”将成为现实。达尔文的理论表明，物种的基因一定能描绘一幅祖先的环境的图像，因为它们就是从那样的环境幸存下来的。物种的基因库是自然选择塑造的泥胚。我在《解开彩虹》里说过：[75]

> 沙漠的风塑造出形态迷人的断崖，海浪拍打出奇异的岸礁，骆驼的DNA从古老的沙漠甚至更古老的海洋里雕琢出来，然后成为现代的骆驼。假如我们能听懂骆驼DNA的语言，它会向我们诉说它祖先变化的世界；假如我们懂得那些语言，金枪鱼和海星的DNA会把“海洋”写进课本，而鼹鼠和蚯蚓的DNA会吐露“地下”的秘密。

我相信，到2050年的时候，我们能读懂那些语言。我们把未知动物的基因组输入计算机，计算机不仅能重塑动物的形态，还能再现它们祖先（当然是自然选择出来产生那个动物的祖先）生活的世界，包括掠食者和牺牲者、寄生者和宿主；包括它们的巢穴，甚至还有它们的希望和恐惧。

如何更直接地重现我们的祖先呢？像侏罗纪公园那样吗？不幸的是，琥珀突变 [76] 的DNA不太可能完好地保存下来，没有哪个摩尔定律，不论儿子的还是孙子的，能把它带回来。不过也许还是有一些办法（其中许多是做梦也没想过的），我们可以通过它们来利用现代DNA数据库里的大量数据——也许我们在2050年之前就已经有那样的数据库了。黑猩猩基因组计划已经起步了，托小摩尔定律的福，它只需要人类基因组计划的部分时间就能完成。

布雷纳在他那篇千年展望的文章（“新世纪的理论生物学”）的最后， [77] 漫不经心地提出了下面的惊人建议：如果完全了解了黑猩猩的基因组，通过拿它跟人类基因做精细的生物智能对比（两者的DNA密码只有1%的差异），就有可能重构我们共同祖先的基因组。这种所谓“丢失的环节”的动物，生活在500万到800万年前的非洲。如果大家接受了布雷纳的幻想，可能不禁想把这种逻辑推广到所有地方，而我并不看好那种诱惑。丢失的环节的基因组计划（MLGP）完成以后，接下来的一步可能是将那些基因组跟人类的基因组并排起来，一一对比它们的基。如果能发现两者的区别（通过类似于前面胚胎学告诉我们的途径），我们将得到一般意义上的近似的南方古猿（*Australopithecus* ）——“露茜”已经成为它的一个种的形象代表。 [78] 到露茜基因组计划完成的时候，胚胎学也该取得进步——只要把重构的基因组注入人的卵细胞，然后植入女性体内，一个新的露茜就将出现在我们面前。这无疑会引发伦理学的忧虑。

尽管我们担心一个重现的南方古猿是否幸福（这至少是一个天生的伦理学问题，而不是杞人忧天地对什么“玩弄上帝”的担心），但是我可以看到，实验不仅会带来科学的好处，也会产生积极的伦理学影响。今天，我们公然的物种歧视侥幸逃脱了可能的惩罚，因为在我们与黑猩猩之间的进化环节的物种全都灭绝了。在澳大利亚著名道德哲学家辛格（Peter Singer）发起的“猿人计划”中，我在自己的文章里指出，那种偶然的灭绝事件应该足以打破人类生命优于一切其他生命的绝对主义估

价。例如，在关于堕胎或干细胞研究的争论中，人们所谓的“为了生命”，总是在为人的生命，而没有什么明确合理的理由。如果一个活生生的露茜出现在我们中间，我们自我陶醉的道德和政治的人类中心论将发生永远的改变。露茜应该“被当作人”吗？这个问题显然很荒谬，就像在南非的法庭上判决一个人是否该“被当作白人”一样。认清那样的荒谬，露茜的再生就没有伦理问题了。

正当伦理学家、道德家和神学家（我想到2050年恐怕还会有神学家）为露茜计划痛苦不堪的时候，相对轻松的生物学家可能已经在忙着一个野心更大的计划：恐龙计划。这个计划，也许可以从帮助鸟长牙开始做起，6000万年了，它们的牙还没长出来。

现代的鸟是从恐龙（至少我们现在喜欢把那些祖先叫恐龙，只要它们像恐龙那样灭绝）演化来的。如果从进化-发育的观点来看现代鸟类和其他幸存的原龙类爬行动物（如鳄鱼）的基因组，到2050年的时候我们也许能重构一般恐龙的基因组。现在，实验室能诱发从小鸡的嘴里生出嫩牙（还能诱发蛇长出腿），这令人鼓舞，说明古老的遗传技术还在发生作用。假如恐龙基因组计划成功了，我们大概可以把基因组植入鸵鸟蛋，让它孵出一只可怕的活蹦乱跳的蜥蜴。尽管有侏罗纪公园，我担心的是自己可能活不到那一天，看不到蜥蜴出来，也不可能把我短小的手臂伸向新生的露茜，满眼泪花地紧握她的大手。

道金斯（**Richard Dawkins**）

道金斯（Richard Dawkins）进化论生物学家，牛津大学“公众理解科学”查尔斯·西蒙尼（Charles Simonyi）讲座教授。他是《自私的基因》（The Selfsh Gene ）、《延伸的表现型》（The Extended Phenotype ）、《失明的钟表匠》（The Blind Watchmaker ）、《流出伊甸园的河》（River out of Eden ）、《攀登绝顶》（Climbing Mount Improbable ）、《解开彩虹》（Unweaving the Rainbow ）等书的作者。他是皇家学会和皇家文学学会的会员，是科学博士，也是荣誉文学博士；他曾入选《牛津引语词典》（Oxford Dictionary of Quotations ），在活着的科学家中这是罕见的。他荣获了1997年国际宇宙奖和2001年Kistler奖。

第二个创生

P.戴维斯

Paul Davies

“火星居住着这样那样的生命，这一点是肯定的，正如不能肯定那是些什么样的生命。”美国天文学家洛维尔（Percival Lowell）用这句戏剧性的话告诉世界，他认为在那颗红色的行星上存在着运河网络。洛维尔猜想火星是一颗垂死的干涸的行星，那里的居民开凿运河来把极地冰盖的融水引向干旱的赤道地区。他画了一幅精细的地图来支持自己的理论。

那是在1906年，火星生命的思想似乎是完全可能的。威尔斯（H.G.Wells）在1898年写的那本很成功的《星际大战》里，[79]尽情发挥了这个想象。许多天文学家至少口头承认火星上可能住着某种形式的生命。后来，在20世纪60年代，发往火星的水手号空间探测器没能揭示那些众说纷纭的运河的任何迹象。1976年，美国国家航空航天局（NASA）的两艘飞船在火星着陆，看到一片荒凉的没有生命的土地。它们挖掘了一些泥土，做了微生物和有机化合物痕迹的分析。结果什么也没发现。那颗红色的星球似乎是沐浴在致命的紫外线中的冻结的荒漠。一句话，火星看起来是死的——太死了。

不过，最近的意见开始改变了。我们说火星不是生命的居所也许太过草率。水手号系列带回的火星表面的早期照片已经显现了干涸的沟渠，而最近几年来自火星全球调查者卫星的更详细的照片，进一步揭示了一些看起来像洪水冲击平原和干涸湖底的地形，甚至还有古海洋的遗迹。显然，火星曾经温暖湿润过，与我们自己的行星没有什么不同。在遥远的过去，那里能繁衍生命吗？今天会不会还有生命生长在某个昏暗的小环境中？

未来50年，我们将有很好的机会来找到这些问题的答案。看样子，新生的天体生物学将在那些年取得巨大进展，而像NASA的起源计划等研究项目有望产生可能的技术，供我们寻找地外生命，回答那个古老的问题：我们是独一无二的吗？作为太阳系中除地球以外人类探索能够到

达的行星，火星将受到特别的关注。它强烈地吸引着我们。那是我们认识第二个创生的唯一机会：在宇宙的另一个地方，生命将从无生命中产生出来。

我们到底该到火星的什么地方去寻找生命呢？它的表面对我们熟悉的任何以液态水为源泉的生命来说都是可怕而恶劣的。两极存在着大量的冰，但那里的温度太低，冰不可能消融。即使消融了，液体也会很快蒸发，因为火星的大气不到地球大气的百分之一。过去，火星一定有很厚的大气层，充满了二氧化碳等温室气体；它能提高温度，为表面长期保持液态水提供足够的压力。据估计，那个“伊甸园”时代在35亿年前就结束了，尽管以后可能还偶尔温暖过。对统治火星表面的生命来说，35亿年太漫长了，所以今天想在那里寻找生命，最大的希望是在表面以下的地带。在过去的20年里，科学家们惊奇地发现有些微生物生活在地壳以下很深的地方。在海底的地下深处也发现了生命。这个深藏的地下生物圈在有些地方延伸了几千米。因为温度随地下的深度而升高，深部的生命很可能是喜欢温暖的，即所谓的嗜热生物。有时候，它们在比水的正常沸点还高的温度下茁壮成长。很多地下生物的能源不是来自阳光，而是来自化学能和热能。有些微生物能利用从地壳渗出的气体和矿物，将其直接转化为生物量，从而维持独立于地面生物的食物链。

地下生命能离开阳光而生存的发现极大增强了我们在火星上找到生命的信心。同地球一样，这颗红色的行星有一个火热的内部，证据是它强烈的火山，有些还在活动。火星的地下无疑存在许多热点，在那里，火山的热量融化了永久的冻结带，为原始生命提供了水源。火星的地下生命也许会通过漏出的气体（如渗到表面的甲烷）而暴露它们的存在。未来号空间探测器将在火星大气中寻找那些来自生物的微小浓度的气体。为了切实研究火星的任何地下微生物，必须深入地下探测。至于该深入多少，没人知道；估计在几米到几千米的范围。计划的火星行动（如欧洲航天局将在2003年6月发射的“贝格尔2号”——一个恰当的名字[80]）将携带钻机和钻头；不过它们似乎深入不到生命幸存的地方。

运回地球分析的表面岩石样本可能提供过去的火星生命的线索。最好的证据应该是发现微生物化石。1996年，NASA科学家宣布，他们发现，在至少1600万年前的彗星撞击中落下的一块火星南极陨石包含着像微生物化石一样的细微特征。那块著名的ALH84001陨石经过了严格的考察，尽管还没有结论，但多数观点认为它不会为火星生命提供决定性

的证据。[81]

NASA正在计划10年内的第一个从火星取回样品的太空行动。一个机器人探测器将从火星表面选择采集一些看起来很有意思的岩石，然后送回地球进行分析。为了绝对安全，岩石将经过严格检疫，不但要保证它们不受地球生命的污染，还要预防任何有毒的火星细菌从它们扩散出去。不过，几乎不可能有什么来自火星的能毁灭全人类的杀手。每个月平均有一块火星陨石落在地球上，在我们的地质历史中，已经有数十亿吨的火星岩石来到地球。有了这样的“陨石交通”，那些样本中的任何火星微生物的祖先可能早就跟着“搭车”来感染过我们了。炮弹轰击和离心实验证明，微生物很容易抵抗陨石从火星喷出时的冲击。一旦进入空间，冰冷的真空条件正好起着保护作用。有些细菌在重压下形成坚固的孢子，能以某种休眠的状态存活很长一段时间。微生物在行星之间穿行的最大灾难来自辐射，但隐藏在直径为几米的岩石中的微生物却可以躲过太阳紫外线、太阳耀斑和所有其他宇宙线（除了极高能的而外）。计算表明，耐寒细菌如果藏在适当的岩石中，可以在环绕太阳的轨道上存活数百万年。陨石经历的最后一场灾难——发生在它们高速进入地球大气的时候——也一定不会出问题，因为摩擦的热量来不及渗透到陨石的内部。总之，似乎没有什么巨大灾难能阻碍可能存在的火星生命成功转移到地球。

与取样计划相关的一些行动也在计划当中，它们更多考虑的是，一旦机器人技术证明是可行的，该在火星的什么地方着陆呢。随着便携式超级计算机、神经网络和精密传感器技术的应用，漫游的飞行探测器会越来越多。过去那种迟缓、笨重的飞行器，如1997年NASA“探险者行动”中的“旅居者”，将被灵巧的漫游者所取代，它能独自探索，能在没有控制的情况下现场选择地形和岩石样品。如果发明一种能在荒漠的表面飞翔滑行的火星飞机，必将极大提高如今受轨道局限的考察技术。

这些技术进步应该能让我们更细致地研究火星表面的物质。不过，寻找包含35亿年化石的火星岩石也不是简单的事情（更不用说活的微生物了）。地球上，只是在很少的地方，而且在经过非常仔细的选择之后，我们才找到过那个年代的化石。无人探测器采集到包含化石的火星岩石的机会是很渺茫的。也许经过几十年的深入研究后，火星是否有过生命，是否如今还有生命的问题，仍然得不到解决。如果那样，解决问题的最后希望就要落在远征火星的宇航员身上了。

把人送上火星不会便宜。送出四个宇航员的价格至少是几百亿美元。不过，削减某些费用还是可能的。独立工程顾问朱布林（Robert Zubrin）在他1996年的《火星问题》一书中指出，火星远征的主要费用是为返航输送燃料。但这也许是不必要的。火星有水和二氧化碳，二者的结合可以产生甲烷，是上好的推进剂。朱布林设想先向火星表面发射一个化学反应器，等它产生了满满一箱燃料才开始远征。宇航员在行星间旅行，来回各需要几个月，会遭遇很多灾难。不过，一旦着陆建立了基地，生命就不至于太危险。正如朱布林说的，火星表面是太阳系里第二个最安全的地方。

在朱布林的计划里，宇航员在火星表面呆两年，在这段时间里，他们要在飞行探测器中进行广泛的研究，寻找生命的迹象。钻探设备可以在宇航员到达之前送去，以便提取来自火星深部的岩石样品。第一个远征队最好在替换他们的队伍到达以后再返回地球，这样，火星就一直有人类的存在。

踏上火星的征程要比阿波罗登上月球艰巨得多。首先需要关注很多技术性的问题。例如，与长期失重相关的医学问题可能会很严峻；利用国际空间站应该能得到一些有价值的经验。尽管计划需要几十年的时间，我几乎找不到什么理由说我们的男女健儿不能在2050年的时候到达火星。

如果我们确实在火星发现了生命，会怎样呢？发现的意义将主要取决于火星生命是否跟地球生命一样。考虑到火星与地球可能相互“交叉感染”，这是很重要的。大量的“陨石飞船”不仅从火星飞到地球，也有从地球飞到火星的（因为地球也偶尔遭到大陨石和彗星的撞击）——当然，由于地球的引力陷阱更深，落向火星的岩石会少一些。微生物可能借着这个过程在往来两个方向传播。两个生物圈的混合将使图像变得很复杂。生命完全可能从一颗行星开始，然后在第二个创生出现之前扩散到另一颗行星。尚在争论的一点是，外来的生命能否很快霸占所有可能的生境和食物资源，从而扼杀第二个创生；或者，两个不同的生物系统能否共同存在于同一颗行星？

火星似乎是更适于生命开始的行星。它比地球小，冷却更快，大概在44亿年前就具备了生命的条件。相反，地球到39亿年前才可能适于生命的栖息。在太阳系45亿年前形成以后，火星和地球都经历过至少7亿

年的强大的陨石和彗星的撞击。最大的一次撞击事件也许能使整个行星成为不毛之地，笼罩在白热的3000℃高温的岩石蒸气中。这个全球大熔炉会给千米深的地下带去一阵热浪，地表下隐藏不够深的生物都将被杀死——但生物也不可能藏得太深，因为太深的地方也会太热。于是，在内部的地热与撞击的热浪这两个上下界限之间，应该存在一个适合生命的区域。在火星上，那个适合深部微生物的区域可能会被抢先占据，成为生命安身立命的好地方。

地球上所有的生命都是相互关联的——就是说，它们来自同一个祖先。栖息在我们生物圈的形形色色的物种不过是同一棵生命大树的不同分支。如果生命从火星开始扩散到地球，那么任何灭绝或者残存的火星生命也只不过代表着那棵大树的另一个分支——也许更低级、更古老，但跟地球生命有相同的起源。到2050年，基因排序技术将高度自动化，而设备也都是便携的。这样，我们有可能在火星基地进行必要的研究，省了一系列检疫程序的麻烦。

假如证明火星生命与地球生命相同，火星就不会成为我们急切寻找的第二个生命的例子。也许我们仍然可以说，生命的起源是奇异的偶然事件，在宇宙间是独一无二的。为了解决生命唯一还是普遍的问题，我们需要看得更远一些。太阳系的另一个可能拥有大量液态水的行星是木星的第二颗卫星欧罗巴。它有一个冰壳，冰壳下面可能是液体的海洋，海洋需要的热量来自欧罗巴绕着木星旋转时产生的潮汐摩擦。因为距离太远，欧罗巴不太可能被地球或火星的生物“传染”。遗憾的是，从我们能预见的任何技术来看，即使在未来50年，载人远征欧罗巴也是不可能的。不过，那时可能会向它发射一个无人探测器。困难在于穿透那厚厚的冰层。一种可能的办法是在探测器上装置一个小核反应堆，这样它可以在前进中融化一路的冰壳。然后，放出一艘小潜艇去探索那黑暗的海底。

天体生物学家们普遍认为，不太可能在太阳系发现任何类型的比简单细菌更复杂的地外生命。复杂的生命大概需要很像地球那样的行星：有厚厚的大气，液态的水，臭氧层，实现大气气体（如二氧化碳）循环的板块构造。在即将到来的几十年，从其他恒星系中寻找类似地球的行星将是一个主要课题。那些恒星的距离太遥远，即使再过50年也不可能去探索它们。如果推进技术没有革命，任何到太阳系外的飞行器都需要几千年的时间才能到达目的地；在可以预料的将来，其他地球的找寻有

赖于改进的观测技术。近年来，天文学家已经用大地基线光学望远镜发现了好几十个外太阳系，但是它们的距离太远了，所用的技术不足以灵敏地确定像地球大小的同样在轨道上环绕着恒星的行星。实现这一点还需要巨大的超精确的空间基线光学系统，它能识别出微弱的行星反射的来自它光亮的母恒星的光，然后从光谱的分析中寻找泄露生命秘密的信号，如行星大气中的氧。这是NASA起源计划里的关键课题。

有人提出一个4光学望远镜的系统：让4个望远镜完全同步地飞行，从而形成一个巨大的干涉仪，它能以前所未有的精度分辨出遥远的天体。这个系统——所谓的TPF（“陆地行星发现者”的简称）——可能在2016年进入太阳轨道。如果行星发现者成功了，跟着的还将有“行星图画者”PI——更大的干涉仪，相当于一个360千米大的望远镜！它将为在太阳系外发现的任何类地行星描绘出特写图像，揭示任何表面生命的活动。洛维尔辛苦观测的火星运河网络尽管把我们引入了歧途，但在他150年后的我们却想象在许多光年外的另一个太阳系有着同样的结构，这是多么有趣的想法。

当然，我们也可能非常幸运地在邻近的星系发现复杂的智慧生命。其他恒星系也可能有地球那样的行星，上面的生命还停留在细菌的水平。复杂生命出现在地球也许是靠了我们太阳系的一些非常特殊的性质，例如，我们的行星有一个异常巨大的月亮，它能稳定地球的运动，防止剧烈的气候变化。月亮可能是地球的外层形成的——在太阳系形成时，一颗过路的火星大小的天体撞击地球，结果产生了月亮。这是非常难得发生的偶然事件。木星也帮了大忙。如果不是它清除了周围的彗星，它们会经常地撞向地球，引发物种的灭绝。这样一些环境，加上其他条件，如我们行星的化学组成和太阳的稳定性，说明像地球一样适合生命的行星在银河系中是相当罕见的。

地外生命的找寻停滞在成功的门口。许多事情还有赖于结果，因为在其他地方寻找生命就是在寻找我们自己——在宏大的宇宙蓝图中，我们是谁？我们的空间算什么？假如生命是落在我们这个宇宙小角落里的惊人的化学意外，像我们这样的智慧生命是独一无二的，我们对地球行星的责任就更加重大。假如我们真的发现了第二个创生，它将永远改变我们的科学、宗教和世界观。如果一个宇宙的自然律喜欢生命，那么在那样的宇宙中，生命就是根本的而非偶然的特征；在那样的宇宙中，我们才真的有家的感觉。

戴维斯（**Paul Davies**）

戴维斯（Paul　Davies）理论物理学家，伦敦帝国学院和昆士兰大学访问教授，是下列畅销科普书的作者：《关于时间》（About Time ）、《上帝的头脑》（The Mind of God ）和《第五奇迹：寻找生命的起源》（The Fifth Miracle：The Search for the Origin of Life ）等。戴维斯的研究主要在量子引力和宇宙学领域，但他的兴趣很广，从粒子物理学到天体生物学。他正在研究生源论问题和宇宙对早期生命演化的影响。多年来，他写了，讲了很多关于科学的深层意义的东西，因为这些工作，他在1995年获得了100万美元的Templeton奖[82]。

预测未来

J. H.荷兰

John H. Holland

长期预测是普罗米修斯式的事业：预测（也许还有预测者）的命运很可能是不幸的。不过，我们很难逃避预测的挑战。有的地方本不该去，可我还是决定要去，问题的关键就在于，我发现，如果把心思集中在预测过程本身的可靠性，把预测结果作为某种形式的图解，那么总有一条曲径通向那个目标。

准确预测的最重要的一个因子是细节的水平。有经验的棋手走出几步之后往往就能预测输赢，但他们很少试着去预测终局会是什么样子。在更复杂的水平上，很多生物学家都会一般地预测进化的生命形式是地球类行星的共同特征，但几乎无人明确预测这样的进化生态系统必然会产生灵长目的生命。任何预测大概都存在同样的问题。

做预测的普通方法是考察目前趋势的延伸。利用这个技术，我们预测了各式各样的事物：如未来的收入，不论国民总产值还是个人收入；如人口的变化，不论人种的改变还是瘟疫的流行。在短时间内，这样的预测是有价值的，但对长期来说，趋势可能是错误的引导，除非基础的

过程有很大的“惯性”，如人口的增长和温室气体的形成就是这种情形。

在当前的技术和社会尺度下，即使我们着眼于人口增长和温室效应，50年也是一个漫长的时间。而且，在那个时间尺度，普遍的特征都决定于我们现在所谓复杂自适应系统（CAS）的影响。CAS由许多相互作用的部分（叫行为者）组成，各行为者在相互作用中彼此适应（或“互相学习”）。股市和免疫系统是大家熟悉的CAS例子。即使在相对短的时间尺度，CAS也表现出多方面的不可加（非线性）效应：自组织、混沌、分形吸引子、冻结事件、杠杆点……结果，不可能将部分的行为加起来得出总体的趋势。另外，我们只有一点零星的CAS理论。由于没有完整的理论，我们也没有普遍和原则性的方法来决定这些不可加效应的影响。于是，当问题涉及CAS时，预测就伴随着灾难。

虽然有这样的警告，我还是认为有可能揭示某些未来可能发生的事情。这种可能性激发了我们构造模型的能力。以计算机为基础的模型，是我们目前最有力的检验不同历史的工具。模型使我们看到了不同作用序列产生的结果，就像飞行专家用飞行模拟器检验飞机设计的极限技术系统。通过模型做预测有几点特别的优越性：

1）预测基础的假定是明确的，因此别人也可以判断假设的相关性和预测的合理性。实际上，他们可以利用或修正这些假定做出自己的预测，从而丰富整个事业。

2）结构好的模型是模块化的，因此可以通过相关的模块和错误反应来跟踪预测的误差，为改进预测提出修正的建议或者新的模块。

3）为了证明预测的强大适应能力，模型能在不同条件和作用（检验极限技术系统）下反复运行。

尽管模型的这些优点对任何预测努力都很重要，但就本文的目的来说，这种方法有着显著的缺点：做计算机模型是很费时间的事情，有时需要几个月，也可能几十年；天气变化模型就是一个例子。本文准备仓促，不可能彻底地讨论模型方法。不过，我们可以把一些构造单元的粗糙描述堆积起来，然后构造出一个那样的模型。粗糙的描述可以带来视频游戏式的图像，人的直觉通常就同那样的图像一道产生。

在动画片的设计中，确立视频游戏模型的起点是确定哪些东西在缓

慢变化，哪些根本没有变化。不变量的和缓慢变化的量提供了支撑预测的框架。一旦框架搭起来，我们就想把看来容易控制和预测的元素装进去。技术的变化往往比社会的变化更容易预测，尽管有些科幻作家——我想起凡尔纳（Jules Verne）、威尔斯（H.G.Wells）和克拉克（Arthur C.Clarke）[83]——很好尝试过长期的社会预测。凡尔纳1863年的小说《20世纪的巴黎》发生在1960年的巴黎，除了前所未有的政治信仰、战争和边境争端发生了变化，他说中了好多事情。[84]不管怎么说，社会的长期预测很少有成功的。摩尔（Gordon Moore）关于计算机硬件能力每18个月翻一倍的预言对几十年是成立的，但我不知道有谁在摩尔做预言的时候预见了Amazon.com和e—Bay。[85]

所以，我下面从看起来简单的事情说起，讲一些我认为有可能在几个相关领域发生的技术变化：计算机化和机器人技术、生物学、运输和空间探测。然后，我将推测这些变化的社会影响——它们对人口、计划和教育、隐私、医药和新探索时代的影响。

技术框架

在10年或20年的时间尺度上，相对容易的预测是粗线条地描绘计算机及其衍生物（如互联网）的未来。摩尔关于硬件进步的定律仍然成立，而软件似乎还将保持它蜗牛似的步伐——有些要10年或20年才翻1倍。软件的迟缓至少跟硬件的迅速一样重要。虽然我们现在有一定能力做出具有针对一定目标的简单知识的计算机，但要让它们拥有更多的人的能力——如灵活的模式识别能力（在一堆混乱的事物中间识别熟悉的对象）或语言的理解力（如理解一部小说）——我们并不比上个世纪中叶好多少。对于能创造、能通过类比和常识进行推理、能做假设的计算机，我们还只有模糊的轮廓。

我们面对CAS问题时，摩尔定律也帮不了什么忙，因为问题尺度的小变化会导致复杂性的大增长。看一个简单例子。在围棋游戏中，如果19×19的棋盘增加一行和一列，那么，不需要改变任何规则，移动10步的可能走法就会增加5倍！如果说这不过是只有几条规则的游戏的情形，那么CAS的情形呢？像市场和政府那样的系统，最粗糙的模型也可能包含着几十条描述相互作用的“定律”。这样的问题不会像深蓝战胜卡斯帕罗夫的例子那样，通过硬件和黑客的方法来解决。不过人们在日常事物的基础上面对CAS，通常是很在行的。用人工智能（AI）先驱明斯

基的话说[86]，奇怪的不是深蓝能以国际大师的水平下棋，而是人类仅凭寻找“10步”的那么一点本领，就能在那样的水平上挑战计算机。自有记载的历史以来，人类就一直在努力去发现产生思想和意识的机制。多数心理学家现在相信，意识关联着中枢神经系统的神经元活动，但意识与神经活动的关系，我们还几乎一无所知。揭示这种关系已经是众所周知的大难题，我想它在未来50年也不会突然就“被解决了”。

社会框架

做社会预测时，我们至少已经知道，人类本性在过去的几千年里，即使有变化，也是变化很慢的。古罗马的元老们从政府信使那儿得到警告，迦太基人的谷物（罗马的“面包篮子”）要歉收了。于是他们垄断了谷物市场，聚敛了大量财富。贪婪的本性，因为一个短期预言而表现出贪婪的行为，在后来的两千年里几乎没有改变。

除了不变的人类本性以外，还有些问题来自高度惯性的过程。人口数量和在人口影响下的过程，如温室气体的聚集，就是一个基本的例子。它们的根基经历了一代又一代，不可能一下子扭转过来。影响我们许多社会议程的问题，都不是权宜之计所能解决的，解决它们还需要对当前行为的长期效应做出合理的预测。

我们今天为增进对生物过程的认识而进行的持续而广泛的努力，还可能一直持续下去，因为它关系着巨大的社会和经济利益。然而问题在于，我们具体做什么——开拓什么，不开拓什么——却受着政治、经济和个人思想等多方面的影响。短期效益与长期探索之间的冲突，在这里表现得尤为强烈。

技术的变革

摩尔定律的持续影响，使我们有可能把20世纪末年的一些普通机器做得更小，并把它们结合起来。我们会做出手表大小的兼有全球通讯、视频、计算、动画设计、定位功能并带3维投影显示的笔记本（类似于《星球大战》里R2D2用的投影仪），可以顶在头上，也可以拿在手上。它会变得和手表一样普通；它还有视频游戏式的界面和为用户服务的知识，最后用起来真的跟笔记本电脑一样容易。

为了获得能自动处理自适应复杂系统的计算机，我们还需要像人脑一样灵活和聪明的软件。我想，只有能学习和进步的软件才能做到这一点。因为涉及CAS的都是重要而普遍的问题——从社会现象（如全球贸易中的市区衰退或涨落）到环境问题（如外来物种的入侵或生态系统的可持续性）——我们将越来越重视具有那些能力的软件。即使今天前进的步伐还很缓慢，我们还是会越来越多地应用那些能根据经验修正自身的软件，以满足不同客户的特殊需求。在50年里，我们可能有像训练有素的助手那样的机器人，尽管它们面对意外事件可能手足无措。我想在50年内我们做不出“有意识的”机器人，尽管我确实认为它们最终会出现。

计算机和自动实验仪器极大地增强了我们对生命和活有机体的认识。到21世纪中叶，20世纪末年的许多医学手段——例如用手术、化疗和辐射来治疗癌症——将被看作跟几百年前的放血术一样无效。我们也许能从简单的没有生命的生物化学物质开始，在一只试管里造出生命，它在技术上能应对所有的疾病。我们几乎一定能制造人工免疫系统来对付生命的病毒和电脑的病毒。我们常常低估了与生俱来的免疫系统的威力；它是那么有效地抵御着各种莫名的入侵者，才使我们大多数人能在很长的时间躲过疾病。更令人惊奇的是，我们发现，从体细胞世代来说，这段没有疾病的时间，相当于从中世纪到现在的人类世代。人工免疫系统的诊断能力还可以帮我们克服某些困难，将基因组序列与基础的信号分子的复杂网络——为生物细胞赋予了遗传性和适应性的所谓生物网络——联系起来。

有一个技术领域，陆地运输领域，其变革晚来了很久。20世纪的汽车为发达国家的人带来了巨大的“活力”，使他们不再像奴隶那样被限定在一个地方。但是我们仍然得面对拥塞的“罗马路”。尽管公司的庞大基础设施体系能承受这样的现实，压缩能量的传输还是即将发生巨大的改变。计算机导航和全球定位可能使一个个的运输更加灵活，不必受路径的限制。未来50年里，有50%的机会出现某种安静的海陆空三线运输。尽管这种新型运输需要获得道路的使用权，但能节省大量基础设施的费用，如公路和桥梁的维护费用。

最后，我想我们将从漫长曲折的太空探索道路走出来。40年前，我们就能用X系列和黑鸟飞到太空的边缘。[87] 但我们抛弃了那一切——毁了关键的生产设备，甚至毁了蓝图——为的是在助推火箭上加载荚

仓，这是一条设计优良的路线，然而走到了尽头。不过，因为以下一些原因，我们将走过这条曲折的道路：

1）我们又开始研究推进系统了，如超音速燃烧冲压引擎（SCRAM），它能让我们飞入太空。

2）能自由操纵行星际空间的国家，像15、16世纪那些能跨越公海的国家一样，占据着明显的科学、军事和经济的优势。

3）20世纪后期的天文学，连同声名显赫的哈勃太空望远镜，向我们展示了“外面”有什么奇迹在等着我们。

社会的变化

50年尺度上的第一重要的事情，是把地球人口进一步降到与可再生资源相当的水平。我们某些最严峻的大尺度问题——如食物短缺、森林破坏、全球变暖、能源危机——都可以归结到人口相对于资源的过剩。摩肩接踵的人带来的自然和精神的压力，不是技术措施所能缓和的。所有国家都把这个问题放到头等重要的位置——如中国——所以，我想在未来50年，人口数量会得到很好的控制，而不会出现像黑死病或世界大战那样的灾难。

实际上，我们正更加敏感地对待其他诸如此类的长期性问题，更认真地对待不同的可能和选择。过去，选择的技巧在棋类和战争等游戏中显得特别重要，但内容总是有限的。视频游戏拓展了内容——像“模拟城市”（Sim City）和“文明”等游戏，极大增强了我们对错综复杂的社会政治相互作用的感受力——界面也更加真实，普通玩家能轻松地选择不同的选项。当这股潮流更紧密地与从气候变化到人工智能的各类事物的精妙模拟联系起来时，会有更多的人（多若干个数量级）有原则地按部就班地探究未来的选择。我前面提到的小型化通用设备——我们称它为“计划者”——将加速这个过程，让未来计划融入每一天的状态。有了视频游戏，就不需要程序专家的意见了，而只需要经验的直觉和探索。“计划者”将使后视频游戏时代的人利用适合用户的现实和可控的三维界面来检验寻常行为的结果。也许，我们会拿飞行模拟游戏来“检验”社会和政治决策的“极限系统”。

当然，“计划者”也牵涉到一些社会问题。有一个问题是早就存在的：对走进专业和学科的人来说，计划者是自然的助手；还有些人则不愿或不能走进来。这两类人之间的知识和收入的差距正越拉越大。在发达国家，几乎每个人都将用计划者那样的21世纪中叶的电话机来探究未来的选择；而在其他地方，知识和收入的鸿沟将更加深广。如今，在南部非洲，读书在9年级以上的人还不足15%；他们多数人都不常用电话，能用计划者的人肯定会更少。

还有一个更大的问题，将令我们所有的人都感到苦恼：如何保护个人的隐私？如何摆脱频繁的监视而自由自在地活动？计划者的视频摄像头（videocam）和快捷通讯能力让每个人都成为新闻人物。这有好的一面：过去因为隐蔽和缺乏证据而猖獗的犯罪，如强奸、抢劫、盗窃等，在现场情景能立刻传播出去的未来，将越来越少。另一方面，也更重要的一面，计划者有可能侵犯别人的隐私。相比之下，喜欢报道任何迎合“大众趣味”的事件（如灾难、人的怪癖之类的东西）的媒体大概要黯然失色了。不受侵犯的隐私和自由，是文明民主的核心（“一个人的家就是他的城堡”[88]）。一旦失去这些权利，专制很快就会出现。到21世纪中叶，在技术上有可能跟踪任何一个人的具体活动。我们将处于外在力量的把握中，仿佛中世纪的奴隶，到临近的村庄去还需要主人的同意。因为言论自由有了限制（我们不能在拥挤的剧院大喊“着火了！”），这里的问题是，通过建立一个法律和习俗组成的系统来严格限制政府和个人对别人隐私的侵犯。这个努力——在《1984》的阴影下，[89]100年后能否成功还是个问题。

增强了对生物学的认识，我们比以往任何时候都能更好地控制疾病和伤痛，从它们带来的痛苦中解脱出来。同时，我们也可能更容易发动生物战争，更容易在基因工程中犯错误。不过，我认为防御会与进攻同步，甚至超前。人工免疫系统对自然和人工系统都形成了有力的保护。人工免疫系统有能力发现那些抵御异常抗原的生物分子，再结合药物设计和生产的自动化技术，药品的成本将大为降低——即使是针对罕见疾病的小批量的药品，也是如此，这就像廉价的CD生产能为很小的听众群录制音乐。治疗成本的减少，加上人工免疫系统的诊断能力，最终能把不断增长的医药费用降下来。

最后，我们即将拥有管领行星际空间的能力，它将挑战当年“新世界”的探索，开创一个发现与振奋的新时代。在50年里，我们可能在月

球、火星和环绕木星的轨道上建立基地。这些基地就像15和16世纪欧洲人在新世界建立的第一批营地，会出现源源不断地激发我们想象力和好奇心的奇迹。处在那个“外面的世界”，我们有更多的机会（像SETI那样）接收银河系其他文明的信号。这些观测（假如做了）可能产生的影响，至少跟古希腊著作的重新发现对中世纪欧洲的影响一样巨大。

荷兰（**John H.Holland**）

荷兰（John H.Holland）是安阿伯密歇根大学心理学和计算机科学与工程学教授，圣塔菲研究所外籍教授兼理事会成员。他的主要研究兴趣是复杂自适应系统（自然的和人工的）、认知过程的计算机模型、计算机思想实验模型的构造。他是著名的“遗传算法之父”，国际遗传与进化计算学会理事。最近的书有《突现：从混沌到秩序》（Emergence：From Chaos to Order ）和《隐藏的秩序：适应如何产生复杂》（Hidden Order：How Adaptation Builds Complexity）。

肉体与机器的结合

R.布鲁克斯

Rodney Brooks

至少在最近500年，科学和技术使我们面临一系列的“泛化”（generalizations），它打开了更广大的天地，清除了我们关于自己和我们的世界是独一无二的认识，使我们变得莫名地不安、愤怒甚至暴烈。17世纪初，伽利略凭着近50年的详细观测数据的支持，就地球在天空的地位问题与教会发生了激烈的冲突。虽然在强大的宗教势力面前，伽利略策略地退却了。但不久就清楚了，地球不是宇宙中心的独一无二的天体，而是围绕着太阳的几颗行星中的一颗。当然，后来发现太阳也只是众多恒星之一；再后来又发现，我们的银河系也不过是一个普通的星系。今天我们的智力难题是，这些令我们越发渺小的发现，是否会落在我们的宇宙本身。

达尔文在他的年代将人类泛化为动物王国的一部分，直接通过血统与它发生联系——这一点甚至在今天美国的理性荒漠还是政治迫害的缘由。20世纪为这个思想增添了一点修饰，根据克里克和沃森的研究，我们发现，我们的许多最基本的基因与酵母菌或果蝇的基因没有多少不同。世纪末的时候，我们又迎来两个“泛化”：我们的生命也许不是源于地球，而是来自其他行星的生命种子。最后我们发现，人类的基因没有想象的那么多，实际上比其他许多动物甚至西红柿还少。不论在哪一点，我们都不是独一无二的。

每一个泛化都曾挑战我们的自我认识。我们已经不那么特殊了，只是更大实在的一个部分。失去特殊的地位往往是难以忍受的，但我们还是慢慢让自己适应了每一个发现带给我们的新世界观。没有哪个发现是突兀而来的。地外智能的发现，假如发现了，也许像一个突然的理性震撼，而且在某种意义上是那样的；不过即使在这里，我们也将逐渐敏感起来，因为我们越来越多的人对SETI（寻找地外智能）有了足够的认识，愿意拿我们还不太习惯的电脑来投入这场战斗。所有这些物种的发现都经过了很多先前的发现、论证和讨论才降临我们。高潮可能是戏剧性的，而前兆也总是存在的。

今天，在这21世纪的开端，我们可以看到未来50年将要发生的另一个泛化的前兆。我们的人性本身将受到威胁，它还可能将我们引向激烈的斗争：从根本上说，什么是理性的思想？什么是宗教的思想？斗争的小冲突已经发生了，没有一点儿动人的地方。我们面临的泛化是，我们人类也成了机器——一样受我们平常用于机器的那些技术的摆布。问题可能更复杂的是，我们技术的基础结构会像最近50年那样发生彻底的改变，而我们身体的技术和生产的技术将泛化为同样的事物。

现代分子生物学的一个尚未阐明的中心原则是，关于生命系统（包括我们自己）的一切事物都是分子相互作用的产物。现代生物学是建立在严格的唯物论基础上的。除了依照各种不同形式的力发生相互作用、服从温度和量子效应引发的随机性的分子以外，再没有别的东西了——没有长生不老的药，没有生命力，没有离开物质基础的思维，也没有灵魂。这些看法不是科学家要讨论的问题，正如没人讨论我们与马铃薯从一个共同的祖先演化而来。假如两个原则——我们生命的分子基础或者生物系统经历过进化的思想——中间有一个是错误的，那么我们的农业、我们的医学、我们的化学工业、我们的制药行业、我们的流行病

学，还有我们的保护工作，都将建立在错误假定的基础上，只能靠它们的运气来工作。关于生命系统还有一些细节需要揭示，在未来的一二十年里，一定会出现几步更大的思想跳跃，也一定能听到不和谐的声音。它们也许会像量子力学对物理、计算机对数学那样，给生物学带来混乱，不过我们目前的认识也不会一股脑地完全被抛弃。中心的原则仍然成立——我们是亿万无意识分子相互作用的产物，不可能是别的：不是燃素或以太，而是日日月月、年年岁岁的千百次新实验所证明的事实。

多数人对过去50年分子生物学的结果一无所知，也一直过得很幸福，不过他们现在也开始来关心了。不久以前，美国总统在全国电视讲话里宣布基于伦理和政治的考虑，政府应该资助什么类型的干细胞研究时，还仔细分析了生物学研究的难以捉摸的细微差别。这肯定不会是我们最后一次看到我们的总统们不知所措的样子；也不会是我们最后一次看到两边的鼓吹者们的传统路线令人眼花缭乱地交叉在一起。当然，我们一定还会看到越来越多的游行示威，有的可能还是暴力的——不单是反对基因食品（我们现在已经有了），还反对那些贬低我们、把我们与我们操纵的人造物等同起来的技术。

我们已经开始将过去50年发展的分子生物学的分析工具转化为工程工具。凭着那些工具，我们正在认识我们有多大能力在生命活动的最基础的水平上把握生命本身——特别是人的生命。

50年前，第二次世界大战刚结束的时候，有过一次工程的转变。在那之前，工程一直是工艺练习，但是大约从1950年起，工程变成了以物理学为基础的一个学科。现在我们看到又一个工程转变正在发生，这次它将变成主要以生物学为基础的学科，尽管也不会失去物理学背景的严格。在MIT人工智能（AI）实验室（我是主任），我每天都能看见这种转变的迹象。我们清理出过去做硅芯片的房间，在那里搭起水下实验室。我们将程序编入结合成基因组的DNA序列，目的是培养出细菌机器人。我们的30年目标是精密地控制生命系统的遗传，这样，我们可以不经过种树，砍树，然后造桌子，而是最终让桌子自己“长”出来。我们把用硅、钢做机器人的实验室改造成了用硅、钢和活细胞做机器人的实验室。我们培养肌肉细胞，用它们作为这些简单机器里的驱动者，例如无缝安装在残疾人身体上的假肢。AI实验室的一些研究如何让机器学习的人员，已经停止建立更好的网络搜索引擎，他们开始设计一些程序，通过学习人类基因组的相互关系来预言疾病的遗传起因。我们把存放机械

CAD（计算机辅助设计）系统的空间改造成了测度人类大脑动力控制的空间，这样，我们最终可能为大脑疾病患者做神经修复。我们的视觉研究者们，冷战时期常为探测俄罗斯坦克建立算法，现在他们建立了专门的视觉系统来指导神经手术。同样的转变正发生在所有的工程部门，不仅在MIT，而且在全世界。

眼前的第一代转变正迅速促进着硅钢技术进入我们的身体。早先从事这些实践的人受现实临床需要的驱使，他们努力去修复伤残退化的躯体。我们很早就有了起搏器和人造臀，最近又有了人造心脏。但是今天更复杂的神经修复术正在成为寻常事情。数万严重失聪的人在他们的耳蜗移植了永久的助听装置；它们提供了6个频率波段，直接刺激耳蜗里某个位置的神经——健康的耳朵在那个位置对那个频率是很敏感的。这些人通过电流对外围神经的直接刺激来听——更具体说，通过硅和“湿”的神经回路的组合来听。

如果类似的视觉移植也同样有效和理想，视网膜黄斑变性的患者将成为最大的受益者。世界各地的研究小组正在考虑的问题是，把硅摄像机芯片植入人的视网膜，然后，要么直接把图像元素与视网膜神经联结起来，要么通过有线或无线将图像发送到以前的大脑后部的视觉过程区域。这种器官的短期植入实验已经做过一系列了，写作本文时，三个患者的视网膜移植已经过了一年多——尽管到现在还没有任何结果发表出来。从技术上说，成功的视觉移植比耳蜗移植更难实现，原因还远不仅仅在于它需要在准确的位置做好摄像机芯片与神经之间的几千个联络（在传送寻常讲话时，只要几组联络就够了）。然而，我们有很多理由相信，视网膜移植最终会像耳蜗移植那样寻常。

几个四肢瘫痪的病人，脊椎骨甚至脑干都严重损伤了，不能说话，不能控制呼吸，需要呼吸器。现在，通过大脑的神经移植，他们只靠思想就能指挥电脑鼠标。通过这种方式，他们又能与外面的世界交流了，并对它进行一定的控制。至少，他们能选择他们想在电脑屏幕上看到的东西，还能（当然多少有些费力）下命令、写信、发电子邮件。在某些实验中，他们能控制在日常生活中帮助他们的机器人。我们可以设想，这些把某些基本的人类尊严还给严重残障者的机器，将继续发展下去，它们的适用范围和适应能力还将与时俱进。

还有很多别的实验——例如，系统训练被打伤的和脊髓损伤者的肌

肉，重新连通帕金森症等类似疾病患者的神经信号——也把硅和钢带进了病人的身体。所有这些实验为我们重新找回大脑关键区域的适应能力带来了希望。

不久以后，这些临床手段就会有选择地应用开来。在未来的10年或20年里，文化将发生转变，为了提高我们在现实世界的实践和认识能力，我们会把机器人技术、把硅和钢用在我们的身体上。眼睛健康的人都可以选择一个对红外或紫外线敏感的器件，安装在自己的某一只眼睛里。我们也许还能直接在大脑安装无线连通的互联网——当然，我们还不知道用它浏览的网页，“看起来”或“感觉起来”会是什么样子。

接着，大约21世纪走过四分之一的时候，我们可能迎来更多的增强我们生物本能的技术。在那个时间框架下，真正的大范围的基因工程——超越我们当前在农业和医药领域里进行的开发——将普遍流行起来。它将应用于石油工业，塑料和其他材料的生产，废物的回收利用，电池的生产，可再生能源以及许多我们现在难以想象的领域。到2025年，我们还能在精确的控制下满怀信心地将这些技术应用于我们的身体。这不是什么奇怪的巧合——同样的科学和技术已经在其他方向上成功应用开了。

早期的某些强化我们生物本能的技术，可能需要在大脑皮层增加神经元的数量。我们正在老鼠身上做这类实验。在处于某个关键生长时期的老鼠的大脑中植入额外的几层神经细胞，它会变得比没有经过细胞移植的同伴更聪明。我们更熟悉在儿童时期控制我们大脑生长的荷尔蒙平衡，也许我们可以在成年的大脑中添加一点神经元，提高几分智商（IQ），恢复我们小时候的记忆能力。疯狂地强化人的能力，很可能产生一些问题和麻烦，但是不要忘了，技术——尽管忽冷忽热，总要继续的。

到21世纪的中点，我们将拥有许许多多新的生物能力。有些在今天看来还是幻想，正如今天的计算机的速度、存储和价格，对1950年工作在第一代数字计算机面前的工程师们来说，也像梦幻一般。我们似乎有理由相信，到2050年的时候，我们不但能在受精时刻干预和选择婴儿的性别，还能选择许多体貌、精神和性格的特征，这可不是寻常的事情。我们已经看到，仅仅决定胎儿性别这一点，在某些地方导致了多么严重的性别比例失调；我们可以预料，新的能力将对世界人口的组成产生深

远的、根本难以预料（在现阶段）的影响。

我们还将有能力改变现有的身体。20世纪末，手术整形和生化美容（例如，用肉毒毒素）在西方世界已经很普遍了；可以预料，50年后我们能看到通过遗传来改变人的身体。这些改变当然是为了延长生命，但也有许多是为了娱乐和生活格调。“人体画廊”将以我们今天难以想象的方式展开。

为改变我们的身体而发展起来的技术也可以用于我们的工业系统。我们现在生产的许多东西，将来可以通过利用基因工程的生命来培育——那些生命在我们的数字控制下完成分子的操作。我们的身体和我们工厂里的材料将成为同样的东西。也许我们还能在思想里把它们分开，就像我们现在能在精神上区别我们的环境与我们在养鸡场饲养的鸡的环境。但是，当那些思想的阴影令我们反思自身有限的存在时，我们把自己作为一个物种的观念将发生改变；我们会发现自己不过是生产系统的一个基本部分而已。

在所有这些科学和技术前进的时候，我们将一次又一次地面对一堆同样令人困惑的问题：什么才是有生命的？什么把事物变成了“人”？什么把事物变成了“亚人”？什么是“超人”？我们能接受什么样的人性的改变？操纵人的生命符合伦理吗？甚至，以特殊的“正确”方式操纵人的生命符合伦理吗？谁规定的“正确”？谁规定的“生命”和“人”？一个科学家对他操纵——或者创造——的形形色色的生命负有什么责任？

这些问题，不仅会在科学范围内被善意地提出，也将在更大的社会里同更广泛的问题一起讨论——从汪达尔行为到 [90] 恐怖主义到最终的战争。

我们的前辈——500年来最早感到不安的那些人——只改变了我们对自己在天地间的地位的认识。在未来50年，新的泛化将赋予我们力量去改变那地位本身。我们正在摆脱自己作为生命和万物秩序的被动观察者的角色，而成为生命和秩序的操纵者。我们将不再看到自己局限于达尔文的进化。现在我们能选择以明确的方式，作为个体、也作为一个物种，参与那场进化。相比之下，我们的核裂变历险不过是一场儿戏。我们需要把我们的骄傲小心地收敛起来——假如我们想让我们的子孙后代能在未来的某一天成为我们星系其他某个地方的SETI的愉快的发现目

标。

布鲁克斯（**Rodney Brooks**）

布鲁克斯（Rodney Brooks）是麻省理工学院（MIT）人工智能实验室主任和计算机科学Fujitsu教授。他还是智能机器人（iRobot）公司（与其他玩具、石油、消费和国防工业等公司合股的一家公司）主席和首席技术官。在莫里斯（Errol Morris）1997年的电影《快捷、便宜、失控》（Fast, Cheap, and Out of Control ）中，布鲁克斯博士是四个主角之一——电影的名字来自他在《英国行星际学会杂志》（Journal of the British Interplanetary Society ）的一篇文章。他是《模型计算机显示》（Model-Based Computer Vision ）、《通用LISP程序》（Programming in Common LISP ）、《寒武纪智能》（Cambrian Intelligence ）、《肉体和机器：机器人如何改变我们》（Flesh and Machines: How Robots Will Change Us ）等书的作者。

物质的未来

P.阿特金

Peter Atkins

化学家是物质的魔术师。他们从泥土、空气和海洋“旋出”新的材料，产生也许在宇宙任何其他地方都不存在的物质形式。不过，跟魔术师不同的是，他们的活动是理性的；他们变幻的基础在于他们深入认识了原子如何联结、如何形成新的化合物。他们对物质的认识，给他们带来力量的认识，从18、19世纪的实验中涌现出来，然后在20世纪随着量子力学在化学的应用而成为定量的知识。在21世纪开始的时候，化学家完全把握了物质。

化学在两个方向上“旋转”，转出新的产品，也转出新的课题。过去50年转出的新产品都在我们的身边：使生命更长久，更愉快和使死亡更少痛苦的药物；使日常生活更活泼的纺织品和染料；使结构更轻更强，

形态更新更奇妙的取代木材和钢铁的塑料和陶瓷；改变了社会的半导体和将再次改变世界的超导体；满足我们生活意愿的燃料。同时，化学也产生了新的课题，这些课题有着各自独立的名称，但本质还是化学的。材料科学是化学——关于产生有特定力学、电学和磁学性质的新材料的化学；分子生物学也是化学——这个20世纪的非凡产物和21世纪生物学和医学的基础，则是涉及与生命相关的可怕的复杂分子的化学；现代医学，除了砍、切、锯那些手术，也在运用着化学，过去的成功凭一点点运气，但现在越来越多地凭认识。跟物质的性质和转换打交道的任何事物，根本说来都是化学的，不论那物质是死的还是活的。

未来50年里，化学家将增强掌握原子并以新模式结合原子的能力。有三条路可走。一条是苦心经营经典的化学技术，例如各种复杂形式的搅拌、加热和混合——它们从炼丹术涌现出来，在我们的实验室已经达到了很精细的程度。第二条是，跟碳的化合物打交道的有机化学家，已经总结了大量以特定形式结合原子的知识，这些知识无疑还会进入智能系统的领域。第三条是，化学合成策略将更多地通过计算机来设计，计算机能利用神经网络选择最佳的前进路线。当化学家想合成更复杂的结构——不单是蛋白质和核酸，还有为了计算和数据存储的有机材料——计算机辅助的化学将发挥越来越基本的作用。我们将在未来50年看到，计算机设计的合成路线将实现后代计算机所需要的复杂材料。计算机肯定会变小。它们必然会及时地用当时可能的最小材料来做；就是说，它们一定用分子来做，因为比分子更小的东西不具备实现一定结构的复杂形式。所以，化学家将以他们制造不太复杂的分子的技艺来制造分子计算机。

有机化学源于化合物的研究。过去认为那些化合物只能从活的有机体构造出来。这种有机化合物只能天然形成的观点，在化学历史的初期就被推翻了；有机化学的非凡活力就是一个明证。结果，有机化学家看到了从头开始合成生命的希望。为做到这一点，他们需要合成DNA或者相当的分子数据存储系统，把它整个地装进人工受精卵（合子）系统。合子系统在合成膜的保护下，配备着为DNA复制提供能量的新陈代谢合成系统。起初，他们还用合成与天然成分的混合物——注入天然卵的一段人工DNA——不过，一个完整人工系统所需要的大多数组成物质，他们已经能够合成了，在未来50年里，他们将制造出大量可以存活的合成蛋白质。我不指望在50年里能从目前的有机化学一下子产生出真正的生命化学——关于活体生命及其性质的合成的化学。但在那段时间里，我

们将看到工作蛋白质产物和很好的近似细胞膜的合成。到21世纪中叶，一个完全的合成生命的零碎都会出现，未来的事情是把它们组织起来。更长远看，没必要老是困在碳上，让硅和锗至少部分进入活体，产生一种全新生命的梦想，也可能成为现实。这些事情的成功，无疑（也理所当然）会引发深远的伦理学问题；不过任何成功的前景都离我们太远，我们暂时还不必为这些问题担心。

尽管有机化学摆脱了有机体是化合物的基本来源的观点，但有些分子还是太复杂了，确实只能由活的有机体产生。化学家将用生命来产生这样的材料，并为它们的产物有效地培育那些生命；所谓的“农业化学”将表现出新的意义。我们正在通过这种方式获取细菌，随着基因工程更加有效，它也会更加重要。当然，我们也没有理由不为更简单的分子（如汽车用的碳水化合物和石化工业的原产品）培育细菌或植物。在未来50年，当我们耗尽了存储的碳水化合物，不得不生产矿物燃料的代用品时，那些应用将变得举足轻重。

除了经典技术和微生物培养，合成的第三个方法将依赖于化学家对单个原子的操纵能力。我们已经可以随意将表面的原子转移到预先选定的位置，将来的分子可以由一个一个的原子构造出来。新生的个别结构可能抵抗不了空气或溶液里经常存在的巨大旋涡。很难想象这些技术能投入生产，但是工业正在变得越来越复杂，所以我们不能排斥这些特制的分子。

特制分子的概念魔幻般展现出一片广阔的纳米工艺和纳米技术的前景。化学无疑将促成小东西的制作，能做出我们现在不得不（实际也是）从大块物质切割下来的小东西。分子工程已经能生产单个分子大小的类似机械零件的东西了，而且这些器械的复杂性还会提高——总有一天，化学家们可以用纳米工艺制造出工程师的所有随身零件，如齿轮、轴承、皮带、夹板等。最早传播纳米技术福音的人认为，化学家能用这些微观零件来构造宏观的机械。例如，一个装着分子制动圆盘的轮子在嵌着分子滚珠的分子轴上旋转。大多数这类想象都假定，原子的性质在很大程度上可以忽略或者任由我们来改造。这样，分子间的力可以减弱，也可以加强，结合的趋势可以忽略，等等。更可能的是，化学家将开发替代宏观机械的分子，它们将考虑而不会忽略组成原子的实际性质。一种诱人的可能是，细菌能通过基因工程“排泄出”整个齿轮、活塞和弹簧，甚至整台机器，排泄物不一定都是有机的，也包含其他成分。

现在，人们对利用碳和氮化硼的纳米管的工艺寄予了很高的希望（但没有多少成功）。为了产生直径为一个原子大小的绝缘线圈，原子弹簧已经融入了碳纳米管——那样的线圈能促进计算机融入极端的纳米结构。我们尽可以期待，专门的计算机能微缩到一粒尘埃，能像浮尘那样喷洒。毕竟，蚂蚁的大脑不比它大多少，却也能表现令人惊奇的特殊行为。

碳纳米管在宏观结构（如吊桥和圆顶）中也可能发挥巨大作用。它们能产生与重量相当的强度。也许难以相信，以纯粹的碳纳米管支撑起来的、外面覆盖纯金刚石片的大地圆顶，有一天能成为保护我们躲避自己的生态罪孽的栖息地；成为让沙漠复苏的乐土；成为火星或行星际空间的根据地。未来的50年，也许正好是碳纳米管成为工业产品的时期。

然而，我们必须承认，未来50年可能不会出现对化学本身的新认识。这门学科已经高度成熟了，似乎不会再出现很多与它的基本原理相关的奇迹。这并不是说化学已经有了可靠的预言能力。20世纪末最大的惊奇之一是发现了富勒烯——外形像足球一样的碳-60分子及其类似的东西[91]，包括碳纳米管——尽管它们在预料之中，却没人把那预言当真过。理论化学惯于以量子理论和统计力学的方法来使观测更加合理，却不大擅长预言。于是，我们可以期待更多的惊奇，不过我们也可以相信，所有那些发现都将在我们现有的认识原则以内。

这不是说理论研究没有意义或者完全脱离实际。计算机在化学的应用已经很重要了，在未来50年里肯定会更加重要。当它们的知识基础更加强大，当化学家更多地利用神经网络来指导它们，当根据个别分子结构计算的总体性质更加可靠，计算机也将日益成为可以信赖的顾问。当前的主要应用是通过计算分子的性质、评估那些性质的潜力，从而鉴别化合物的药理学行为。原则上讲，这种鉴别可以成年地缩短药物的开发时间。现在全世界都在进行那样的鉴别计算，它（像当前的地外智能探索那样）利用全球的联网计算机的间歇时间扫描可能具有药理学活性的分子。这类计算机应用无疑还会增长，特别是人类基因组计划的最终完成为它提供了大量的数据。

计算机会更多地应用于化学，指导其他化合物的合成，其中也包括催化剂——它能不损自身地刺激特定反应以可观的速率进行。（在中文里，“催化”也有“媒”的意思，很好把握了它的实质。）催化剂是工业的

荷尔蒙，离开了它们，化学工业也就不复存在了。多数石化工业的主要研究是发现和开发更高效、便宜、持久和更多选择性的催化剂。只要有化工产品的生产，就离不开催化剂的应用。早期的催化剂，如铁块或铂和铑制成的丝网，显然是非常简单的，但现在的催化剂正变得更复杂。在未来50年，化学家将开发固体催化剂和一系列能溶解在液体中、通过溶液发挥作用的新的均一催化剂。固体催化剂将更多地采用多微孔材料，它们充满了分子大小的迷宫似的孔洞、隧道和网格。多孔材料的好处是具有很大的表面积（它们几乎整体都是表面），对渗透它们的分子的类型和大小有高度的选择性。计算机正更多地用来发现这些材料的功能并重新设计它们。未来50年将涌现出合理设计和应用这些材料的潮流，大量新的廉价材料将从应用它们的工业中产生出来。

我到现在还没说化学活动的也许更传统的方面——通过材料的分析来认识出现的物质。过去50年化学分析的进步几乎完全归功于三个方法的发展：首先是色谱法，它让物质以不同的速度通过细长管道；其次是质量光谱测定法，分子在其中分离，然后通过它们产生的碎片来推测原来的性质。这两个技术都很灵敏而且经常联合运用，当然还需要进一步的改进，以识别数量更小的材料。第三是分光镜的一整套技术，它监测不同类型的电磁辐射（红外线、可见光、紫外线、微波等）的吸收。在这些技术中，现在为止最有用的是核磁共振（NMR），它也是医学上有效应用的磁共振图像技术的基础[92]。

事实证明，NMR是化学工具中最具适应能力的。大约50年前，在技术发展的初期，它监测的是一个氢原子核在强磁场中掉转方向时吸收无线电波的情况。自那时起，随着相关的电子学日益复杂，我们能使整个氢原子群和其他类型的原子核群集体掉转方向，NMR技术也跟着成熟起来了。我想强调的是，这项技术在几十年里几乎是有机成长起来的，而且显现了未来50年复杂性的每一个增长的信号。它似乎达到了一个成功的高地，只有加入更多的复杂性才能把它重新激发起来。每多一点复杂性，化学家就能多吸取一点关于样品分子的信息；最近的主要成果是确定了蛋白质在接近自然栖息条件（细胞内部的液体环境）下的结构。那个技术的强大适应能力也许还能帮助我们制造量子计算机。谁能想象——也许在未来的某一天，当量子计算实现的时候，一台NMR光谱仪竟然也开始来思考它正在观测的分子！

化学不仅是关于组成的，也同样是关于结构的。化学家寻求通过分

子的形态、大小和原子排列来认识它们的性质。这样，他们能将水的许多性质追溯到水分子是V型的事实；他们寻求通过蛋白质这些重要分子的螺旋、片段、扭曲和转折来认识它们的性质。当然，这里也存在着我们预期将在未来50年解决的理论和实验的问题。

当今吸引众多注意的一个理论问题是这样的：给定氨基酸的序列，它所形成的多肽链（即蛋白质分子的骨架链）在其自然环境下会呈现什么形态？这是分子生物学的一个关键问题，因为蛋白质分子的形态有效决定着它的功能。即使不考虑从组成追溯功能所获得的纯粹而非常有趣的知识，功能的确立也可以认为是人类基因组计划的一个基本组成部分。在这里，我们从DNA所编码的蛋白质出发，通过它们的组成和形态来追踪它们行使的功能，然后认识其中的信息。对这个所谓“蛋白质折叠问题”，一个解决方法是计算，不过需要强大的计算机来分析长长的多肽链可能产生并束缚在其中的扭曲结构。这个问题正在逐渐展开，在未来几十年可能占据大块的化学，吸引众多的化学家。如果多肽链没有为它要发挥的功能选择正确的形态，合成多肽链——那是非常简单的事情——就毫无意义了。

决定形态的实验问题，在很大程度上通过X射线衍射的引进而解决了。那个技术已经100年了。20世纪中叶，它因为确定了DNA和许多重要蛋白质（如溶菌酶、胰岛素和血色素）的结构而达到了今天那么神圣的地位。技术最新的发展在于极高强度的X射线源的应用，有望成为未来几十年发展的基础。那些射线来自同步加速器——一种巨大的环形机器，约束其中的电子以很高的速度做圆周运动，在改变方向时发出X射线。同步加速器是国家级的大型设备，应用在世界各地的很多研究中心；它们产生的高强度X射线能使我们更快、更清晰地获得X射线衍射图样。我们将逐渐能确定溶液中的分子的结构，甚至观察正在发生的反应。

合成、分析和结构，是化学的三个主要组成部分，我们现在都讲过了。最后我们来说化学反应，即一种物质变成另一种物质的实际过程。最近，因为脉冲激光的应用，分光镜技术取得了新的进步，使化学家能检验飞秒（10^{-15} 秒，一千万亿分之一秒）尺度的反应事件。在那样的时间尺度，飞行中的原子也几乎没有运动。迄今为止，只有极简单的反应经过那样小时标的检验，但是可以想象，技术的发展使我们能以这种方式去检验实际发生的反应，甚至那些酶催化的反应。这样，我们可以

获得反应过程的一帧帧电影画面，注视那些处在邻近瞬间的原子和分子，最后使我们真正深入地认识在我们魔幻技术操纵下的物质是如何形成的。

阿特金（**Peter Atkins**）

阿特金（Peter Atkins）是牛津大学化学教授和林肯学院会员。他的研究领域是理论化学，特别是磁共振和分子的电磁性质。近年来，他几乎把所有时间都投入了写作，包括教科书：《普通化学》《物理化学》《无机化学》《分子量子力学》《量子》《物理化学概念》；普及读物：《分子》（Molecules）、《第二定律》（The Second Law ）、《原子、电子和变化》（Atoms, Electrons, and Change ）以及最近的《周期王国》（The Periodic Kingdom ）。

我们会更聪明吗

R. C.香克

Roger C. Schank

智慧是绝对的吗？人类会随时间流逝而变得更聪明吗？当然，这要看我们所说的智慧是什么意思。我们肯定在获得更多的知识，或者至少看起来是那样的。然而，尽管今天的普通小孩接触的知识远远超过了50年前的小孩，可还是有人呼吁，我们的孩子今天受的教育不如50年前了，学校令我们失望了。

今天，智慧意味着什么，受教育意味着什么，这些问题不是我们科学探索的核心问题，也不是普通话题的核心问题。不过，我们的生活却依赖于我们对智慧和教育的隐约认识。这些观念将在未来50年面临严峻的挑战。

大约10年前，我应邀参加《不列颠百科全书》编辑委员会。另一些委员多数是80多岁的老人和人文学者。因为我是科学家，又比他们每个人都年轻很多，所以我说的多数事情都会招来他们奇异的眼光。我问他

们是否愿意做一部比现在大10倍而所需费用相同的百科全书？他们回答说不，现在的百科全书的信息量正合适。我说，如果他们真的那么想，不出10年就会失业。他们一点儿也不明白我的意思——尽管我费力解释过所谓“www”来临了。在后来的一个会上，听了我对未来的同样主张后，20世纪40年代的一位文学大家法迪曼（Clifton Fadiman）回答说，“我想我们都不得不接受这样一个事实，我们所受教育的头脑将不如很快来掌管百科全书这样的机构了。”

那时的《不列颠百科全书》董事会主席是已故的阿德勒（Mortimer Adler），他还负责一套《西方世界经典》丛书[93]，现在还在整套地卖。这些书代表了世界智慧的伟大著作——至少阿德勒和他的同事们是这样想的——而包括在系列里的书几乎都是在20世纪以前写的。我问阿德勒是否认为还有些新书可以添加进来，他说大多数重要的思想都已经写出来了。

所有伟大的思想已经有人想过了，这个观点长期以来一直流行在教育和智慧的美式观念中。下面是哈佛学院1745年的入学条件：

凡能即席（*ex temporare*）阅读塔利或其他经典拉丁作者，[94]能以本人能力（Suo Marte）写作纯正拉丁诗句和散文，能正确完成希腊语名词和动词所有词尾变化之学者，均可进入本学院；未达此条件者不得提出入学要求。

经典丛书和1745年的哈佛的共同点是，它们在根本上假定人和人的制度的研究在古代已经被充分把握了，所以教育要求我们熟悉和精通前辈们的思想。在这个意义上，受过教育的人就是能渊博地谈论各种历史、哲学和文学问题的人。接受教育——从而获得智慧——在过去的一个世纪和更以前的时代，就是积累事实，学会寻章摘句，学会引经据典。教育意味着知识的积累，智慧在大众的想象中也往往只不过意味着能展示积累的东西。

但是，假如知识从四面墙壁涌来，会发生什么呢？

未来50年中，知识是很容易获得的，只要你能大声地说出想知道的东西，四面的墙壁就能有回应——当然，那墙壁的背后藏着好多的技术。如果一时忘了弗洛伊德关于“超我”都说过什么，那也不要紧，你可

以就近找一台机器，问它弗洛伊德说过什么，然后听弗洛伊德（或某个看起来和听起来很像他的人）说，还能找到五个观点相反的大人物，随时准备提出他们的观点——假如你愿意和他们一起讨论的话。

但是，智慧是获得问题答案的能力还是学会提出问题的能力？当答案失去了意义，问题就更显得重要。我们在依赖答案的社会生活得太久了。那迹象随处可见：人们看的电视节目，如《危险》（*Jeopardy*）[95] 和《谁想当百万富翁？》；玩的游戏，如“追问到底”（Trivial Pursuit）；而在多数学校，答案更是统治着一切。于是，我们的学校越来越只关心考试。学校成了学习答案而不是学会提问的地方。

新技术会改变这一切。计算器出现的时候，人们问是不是还可以用它来数学考试，因为从今往后这种东西总能得到。结果，数学考试开始关注比长除法更实在的问题。人工智能进入日常生活会带来同样的结果。当机器无处不在并能回答与我们有关的任何问题，我们就失去了过去做实际的知识仓库的价值。过去人们认为城里最有学问的人有知识去影响别人，而其余的人被迫老老实实地听他们的话，建立在这个基础上的旧的教育观念，将被新的知识获取观念所代替。知识将不再被看作我们需要的日用品。社会中任何得来容易的东西都会贬值，知识也不例外。

真正有价值的应该是好的问题。我们会听到人们说，计算机只能带你走这么远。

想象一下，你坐在客厅里跟朋友说话，吵了起来。你面对墙壁等待回答，你想知道谁是对的，墙壁告诉你，有很多虚拟的人可以加入你们的谈话。你选择了几个你听说过的或者以前跟你谈过话的人。接着，一场谈话就开始了。最后，计算机收集的知识到了极限的尽头，墙壁也不知道更多的东西。你惊叹道，“真是一个有趣的问题！”知道一个好问题，你就能和对同样问题感兴趣的人一起讨论。你把问题告诉墙壁，那些同样兴趣的人——那些和你一样超越了软件的人们——会突然（虚拟地）降临你的客厅。在可能发生这样的事情的世界里，受教育是什么意思？有智慧是什么意思？

要考虑教育的问题，人们不得不问问那个世界的小孩的生活会是什么样的。未来50年里，我们所了解的那种学校将因为失去意义而萎缩。

虚拟的经历就在身边，世界上最好的老师随时可以来，为什么还去学校学那些事实呢？教育将意味着，从两岁开始就在智慧——能回答问题并能提出新问题——指导下探索有意义的世界。一个个新的天地将向着好奇的孩子敞开。在这样的社会，教育关心的是你走进了什么样的虚拟（后来成为现实）的世界，在那些世界学会了多少事情。

对法迪曼上面的那句话，我的回答是，人们的头脑不会没有好的教育，而是会有不同的教育。在法迪曼的世界里，有教养的人就是那些在哈佛之类的学校受过训练的熟悉西方思想主要观点的人。他的教育观当然不包括能用JAVA编程序或者了解神经科学的基本概念。未来50年里还会有哈佛，但是它的入学要求的价值将发生巨大的改变。

在最深层的意义上，教育总是更多地关心“行”，而不是“知”。几千年来，许多学者都指出过这一点：亚里士多德（“对那些只有在学过以后才能做的事情，我们通过做来学”）、伽利略（“你不可能教人任何事情，而只能帮助他自己去发现”）、尼尔（“听了就忘；看见的还记得；做了才理解”）、爱因斯坦（“知识的唯一源泉是经验”）。不过，学校把这些教导都忘了，而是选择了——借杜威（John Dewey）的话说——“通过灌输来教”。[96]

新生的虚拟学校将取代现在的那些学校，吸引更多的学生，倒不是因为他们有什么承诺，而是因为它们能提供很多经验。因为那些经验是现成的，时刻等着好学的人，多数学生远不到18岁就可以开始读大学了。在不同虚拟经历上的成功将激励我们去迎接新的历险，很像我们今天的电脑游戏。资格审查部门会更加关注你能做什么——你得过哪些虚拟的荣誉奖章——而不在乎你学过哪些课程。

人们将努力创造经验。将来，世界各地的物理学家不会在哈佛或哥伦比亚那样的大学开物理课，他们会和虚拟教育世界的设计者们一道开发能产生物理经验的软件。这些经验向每一个人开放。过去，学校通过考试来检验学生对课程的学习情况，人们于是认为，最聪明的学生就是那些在学校考试中得分最高的，这个观点要变了——最聪明的学生是那些为软件提出了需要大家来解答的问题的人。智慧将意味着达到学习经验的极限的能力。

我们的社会在总体上会因为这些新发现而更聪明吗？从自然的思维

能力看，人们现在与过去或者将来是一样聪明的。不过，一个卓越的穴居者，对世界只有那么一点儿知识，多年来只得到那么有限的智慧，他只能用熟悉的工具劳作。他本可以跟后来的古希腊人一样很好地理解人类的本性，本可以跟后来的古希腊人一样地充满智慧，但在任何绝对的意义上说，他都不够聪明，因为他有太多没有经历过的事情。

当然，我们也可以这样来看古希腊人。亚里士多德是杰出的，因为他处理过我们今天还在处理的问题，而且他对那些问题有过很好的见识。不过，他在处理他毫无经验的问题时也几乎是天真可笑的——而我们对那些问题却有着多得多的经验。每一代人都在为下一代改进那些经验，但在下一代却可能出现飞跃。50年后的人可能会笑话我们今天还有老师、课堂和书本。他们可能问我们为什么经历那么长的时间才改变了教育观念，为什么会把SAT分数看得那么重 [97] ，为什么会在任何情况下都把记住答案作为智慧的标志。18世纪大胆提出的教育是关于国家教化的观念，今天似乎没有人赞同了，将来更显得可怕。政府对知识的控制——在有的国家还很普遍，在没有计算机的国家也可能还存在——将一去不复返了。知识到处都有，而且那么容易获得，没有人能够阻挠任何人去经历任何事情。政府将不得不抛弃像今天这样统治教育的幻想，他们将无力控制四面八方的虚拟世界，就像今天想控制电视和电脑的国家都一个接着一个地失败了。

我们开始懂得，在未来50年，个人的经验和拓展经验的能力是智慧的最终测度，也是自由的最终表现。创造虚拟经验将成为重要产业。我们的家将被虚拟经验占领，我们的学校将被虚拟经验取代。我们今天在电脑游戏和科幻电影看到的将成为我们未来的现实。今天，像“无尽的任务”（everquest）那样的游戏吸引着成千上万的玩家，他们在虚拟的世界里去争取地位，建立社会关系，获得各种虚拟的东西。游戏对参与者来说是那么真实，他们需要的那些虚拟的东西竟可以通过e—Bay来买卖（价格很高的）。很多玩家的社会生活也完全建立在那些游戏的基础上。将来，这些世界会变得更复杂，与真实世界交织得更紧密。

我们真的可以在任何时候到任何我们想去的任何地方，而任何人想问我们的问题总是我们去过什么地方，在那儿有过什么经历。我们将寻找那些在虚拟的世界里比我们更有经验的人。我们将理解，尚未回答的问题和能以批评的眼光来审视那些问题的人，才是真正测度一切智慧的决定性因素。当然，最后这一点在今天的大学里已经普遍认识了，但还

没有在商业和政府中得到真正理解。政治家需要单纯的观点，老师需要正确的答案，商人需要解决的办法，冒险的资本家需要利润，媒体需要本国的肥皂剧，认证机构需要考试的分数。在那样的社会里被认为聪明的人是那些为社会提供了所需要的东西的人。从这样的供求观点看智慧和知识，恐怕法迪曼也会觉得被遗忘了。不过他和他的那一代人还能守住自己，超越那一切，继续谈论那些伟大的经典。

我曾应邀考察过一些技术学院，看他们是如何教育的。在一个厨师班上，每个学生有一套自己的厨具，正忙着做吃的。我只能说，我说不出什么有意思的话。学校通过让学生做来教他们做。在技术院校，这算不得多激进的观点，但在更高等的院校，它似乎真的很激进。当做事情的工具越来越多时，做才是重要的。在我工作的卡耐基梅隆，新学生一走进校园就必须把自己的计算机装配起来，未来的四年就用那台机子。可以相信，一旦他们为自己做了一台计算机，他们就会认识机器是怎么工作的。

在以现实行为环境为基础的教育体制下，重要的不是我们知道什么，而是我们能做什么。未来关于智慧的主要问题，将围绕着虚拟教育世界中学生相互作用的本质问题而展开。

当教育的环境需要问题、询问如何获得问题、需要知道提出那些问题的经验时，那么计算机带来的深刻变革就实现了。从我们不再畏惧新的经历这一点说，我们都会更聪明——聪明得多。我们将知道如何发现那些经验，我们将从那些经验里成长起来。我们的头脑将经历不同的教育，统治我们智力社会的既不是人文学者，也不是科学家，而是有经验的人——那些生在那个世界并因此而好奇的人。

香克（**Roger C.Schank**）

香克（Roger C.Schank）是一流的人工智能专家，认知人文项目负责人，卡耐基梅隆计算机学院杰出教授。曾任西北大学学习科学研究所所长，是那里的荣誉教授。他的书包括：《动态记忆：在计算机和人群中学习的理论》（Dynamic Memory: A Theory of Learning in Computers and People ）、《给我讲一个故事：真实和人工记忆新视点》（Tell Me a Story: A New Look at Real and Artifcial Memory ）、《专家心理指南》（The Connoiseur’s Guide

to Mind ）、《虚拟学习：培养高技能劳力的革命性新途径》（Virtual Learning：A Revolutionary Approach to Building a Highly Skilled Workforce ）、《线条外的颜色：打破一切规矩，培养聪明孩子》（Coloring Outside the Lines：Raising a Smart Kid by Breaking All the Rules ）、《设计世界E学习课堂》（Designing World Class E-Learning ）。

复杂性的顶点

J.拉尼尔

Jaron Lanier

计算技术流行的第一个50年，大体上跨越了20世纪的后一半。这是忘形的夸张与沮丧的失落疯狂交替的50年。夸张来自计算机的创立者们：图灵想知道，机器，特别是他那抽象的“万能机器”，能否最终成为在精神上与人同等的东西；香农（Claude Shannon）以同样的激情把“信息”定义为具有最广大的范围、囊括了一切热力学过程的东西。

我们同样可以说，因为所有生命都是化学相互作用构成的，所以任何化学仪器都可以作为人的某种原生的形式。我们没有那样说，因为生命物质的化学复杂性与可以在当代化学实验室里研究的复杂性之间存在着显然的差别。我们可以直观感觉这种差别。反过来，我们却不能直觉地分辨不同信息系统之间的复杂性的差别。一个说自己在研究“人工智能”的严肃的知识群体，早在20世纪50年代的时候，就相信计算机很快能流利地说自然的语言。当然，这在今天还没有发生，而且我们仍然感觉不出理解自然语言到底是多大的问题，需要多长的时间才能解决。

夸张还在继续着。在计算机科学部门里，甚至总能找到那样的一群精英，他们相信在未来的50年必然会出现一个“奇点”。奇点出现的时候，计算机已经聪明绝顶、威力无穷了，它不但取代了作为占统治地位的生命形式的人，还主宰了物质和能量。它像神或上帝那样存在，完全超越了人的概念。上面这些句子，我写起来都觉得奇怪，却是我的许多同事的想法的真实写照。

有些读者会注意到我因为同样夸大了“虚拟现实”而受到了批评。[98] 不过常常被误会的是，虚拟现实的目标不是彻底描述或再现物理现实（那几乎是不可能的妄想），而是充分理解人类的认知，把人的神经系统投入到幻想的进化游戏。虚拟现实在本质上是关于舞台魔术的极限的研究，而不是关于约化物理现实的研究。

当人们在理论上堂吉诃德式地夸耀计算机的力量时，真实的信息系统的表现却一个接着一个地令人惭愧和失望。计算机是唯一注定会在正常运行中经常发生意外事故的工业产品。维护信息系统的费用几乎总是被低估了，甚至可以说这是现代商务的惯例。

具体说来，正是软件这个东西，我们不可能照预计的价格来控制它，不过那只是一定类型的软件。硬件会以摩尔定律的指数形式变得更小，更快，更便宜。正是这么飞速的进步点燃了狂热夸张的激情。封闭系统的软件，有着小巧而且可以固定下来的不变的界面，也能做得很可靠，但不会很便宜。这类软件的一个例子是驱动现代飞机（如空中客车）的密码。那类我们似乎还不能控制的软件，有着复杂而且随周围环境改变的界面，例子是个人计算机软件，都知道那是很难控制的。重要的是不要混淆这两类软件。20世纪末广泛流行着一种奇怪的妄想，说千年虫（Y2K bugs）会产生大规模的破坏。破坏没有发生，原因是多数基础软件都是可以控制的，虽然代价很高。

未来50年里，计算机科学中的这两种趋势——能力的夸大和代价的低估——可能还会继续下去。这种景象大概可以叫“帮助台的行星”，生在其中的人们将费巨大的力气来维护十分庞大的软件系统。[99] 这个景象倒不是一点儿吸引力也没有，因为它能让人的力量得到充分的发挥。暗淡的未来是逃避不了的，不过我们还是值得想象一个可能带来崭新东西的计算机科学的新时期。

首先，计算机科学必须回到起点，重新考虑信息与物理过程的关系。香农破天荒地将可以测度的信息量与物理系统的熵联系起来，但这个孤立的公式容易使人误解。实际上，并不是所有信息量都能测度，因为有些信息会比其他信息更重要。物理系统中多数有可能被测度的信息量实际上都在统计分布的海洋中丢失了。20世纪后期有个流行的比喻说，蝴蝶轻轻抖动一下翅膀，几个星期以后也许能在地球的另一端引发大风暴。与这种观念相关的一个问题是，即使它偶尔是对的，也没有那

么多风暴来回应那么多的蝴蝶。[100] 我们也许可以说可测的信息有着不同的“因势” [101] 。香农的信息大概应该重新命名为“势信息”。如果一个信息单元很重要——就是说，它有很高的因势——就一定要研究它；它必然是系统的关键部分。这产生了有时被称作“语义学”的东西，也就是计算能在其中显出意义的那个环境。

在计算机科学里也总是存在一个观测者的问题（只是偶尔被承认）。为了表述这个问题，一种办法是考虑某个对人类语言、历史和文化一无所知的异类生命。那些异类不可能漂浮在星际空间编一本莎士比亚戏剧，同样也不可能可靠地重构一台孤立的个人电脑的意义和功能。

这不是遥远的理论问题，而是急迫的实际问题。由于当前人类工程师具体分析和控制软件的能力限制了软件的复杂性，可以说，我们已经达到了我们所知软件的复杂性的顶点。如果不寻求新的软件思想和设计途径，那么，不论处理器变得有多快、多丰富、多奇特，我们也写不出超过千万行代码的程序。

在20世纪中叶计算机科学的黎明，我们唯一的直观感觉的信息经验是通过线路来发送脉冲。早期形式的信息论——在标准课程里至今仍然占着统治地位——关心的是线路末端世界的一个个采样点。于是，如我们所知，计算机的结构设计就围绕着这些模拟线路。源代码是对脉冲的模拟，它可以像传递的变量或者消息那样连续地向线路发送出去。

为了让一根线上的脉冲有意义，需要一个协议来根据信号的顺序为它赋予某种意义。计算机科学的前半个世纪几乎就是在这样的协议激发下走过的。当然有成功的，如开通互联网的协议。但这不是自然系统的工作方式。虽然在理论上我们可以用20世纪的算术协议来解释视觉皮层从光学神经接收到信号以后的行为，但这种做法会在我们完全不可能把握的尺度上涉及巨大的复杂性。显然，协议的遵守并不能有效解释同时接收大量输入信号的系统，而且可能也不足以设计庞大的系统。如果我们以一个能在多点取样的面来替代线的概念，那么我们必须脱离算术协议而走进一套新的技术，包括模式的分类、绝对确定的预测模型的自动维护。

当前的一个实际问题可以说明这一点：我多年来和外科医生一起做模拟模型，以帮助他们为特定的病人制定医疗方案。以今天的标准看，

模型很复杂。为了保证它们有用，每一个模型都由有多年经验的专家群体建立和维护，而且都必须经过几千个病人的检验。

现在，假定某医学院的一个小组用10年的时间做出了绝妙的虚拟心脏，显示了良好的手术应用前景。同时，在另一个学院，一个类似的小组用10年时间研究了虚拟肺。我们假定，两个小组愿意把他们的成果结合在一个虚拟的胸腔里。

两个小组几乎肯定会运用互不相容的协议。他们不但可能选择不同的基本机器、操作系统、执行语言等，还可能走不同的概念路线。也许一个小组强调维护整体的、自上而下的约束，而另一个小组喜欢自下而上的组织法则；一个小组可能强调对象的语义学，另一个小组则试图逼近一个连续的系统。在当前的技术水平下，两个小组可以就能在它们之间传达的线路信号达成一个协议。这样的协议是有问题的。在这种情形，复杂性也许只能起着某种抑制作用。几年后我们会明白的。目前正在努力。如果可能有器官之间的协议，那么达成一个协议就等于被迫在器官模拟技术上做一次可悲的交易。一个工作协议几乎肯定会损害我们改进任何相关组成器官模拟的前景。

为了理解为什么会这样，我们需要深入信息系统遗留的问题。最能刻画我们时下软件的形容词是“脆弱”，崩溃了也不会屈服。这是过分遵守协议的结果，那原本是一个不可原谅的要求。因为这些根本的脆弱性，软件是一层层搭建起来的，为了发掘那些已经为众多用户以不同方式所依赖的协议，还不知道有多复杂，需要花多少钱。于是我们有了“锁定”现象：有些软件实际上成了强制性的。锁定现象被20世纪末的软件销售者们操纵着，为他们带来了有史以来最大的财富。

除了锁定，软件还有更令人讨厌的特征，我称它为“沉淀”。在软件沉淀的过程中，协议连同嵌在协议里的思想都成了强制性的。文件的思想就是一个例子。大约1984年前，人们还在争论文件是不是好办法。有些计算机科学家觉得共享的信息最好能有细粒化的结构——也就是一个由字母似的基本小单元构成的单独的整体文件。实际上，第一代国际版的Macintosh [102] 计算机并没有使用文件。可是，发行版的Macintosh有了文件，而且Windows、Unix和其他几种广泛应用的系统也都有了文件。现在我们将文件作为像光子一样基本的严峻现实告诉学生，尽管它们是人类发明的。

回头来看虚拟的心和肺。一旦两个工程小组达成了协议，那协议就成为他们的主人，因为他们不得不同时改变自己来修正它；那是一个复杂而昂贵的任务，实际上是不可能的。不论协议达成的时候流行什么样的关于器官联络的思想，它都会“沉淀”下来。思想也就停止了。

因此，未来50年计算机科学的一个美好愿望是为大系统的组成寻求一种新的联络方法，以取代对协议的依赖。在心和肺的情形，那种方法已经隐约显露出来了。

我们设想，每个器官都认为另一个是真实的血肉器官，连接着一个真实的传感器；每个器官可以测量另一个器官里的基本性质，如温度、压力和在时间和空间的某个点的化学组成；每个器官在另一个器官看来都是一个能在不同程度采集样本的表面，但器官之间不存在更高级的参数交流；除了可能的物理测量确定的那些低级协议而外，没有别的协议。

为了实现这个计划，每个小组都需要学会识别另一个小组模拟的模式。心脏不再能发送心跳的消息；它只能由肺通过诸如流体运动和组织位移等过程来推测。每个小组还需要学会建立另一个器官的模型，以帮助解释测量结果。这些模型也许不能作为独立、分离的结构而存在，但可能隐含在所选择的信号过程的方法中，而且几乎肯定能在运用中进行自我调节。

这种模型的建立大概可以叫做“统计表面联结”。假如它对器官模拟有用，对一般的计算机结构也可能有用。也许将来会有某种操作系统，其组成部分能相互识别，解释甚至预测。这样的系统不容易产生灾难性的崩溃。现在还无法知道这种计划能运行多好，不过，假如计算机要超过我们今天知道如何把握的尺寸，很可能需要采纳某些统计联结。

在目前的情况下，我们一贯把信息系统的描述（协议的遵守）看得非常乏味和低下，而且高傲地以纯理论的眼光来俯瞰复杂性。但是我们缺乏一个中间的视点——通过巨大组成部分之间的关系来形式地理解系统的复杂性。如果我们能以外科医生的观点将人体模拟为一张信息表面图，这样的技术有可能推广到认识生命系统的其他问题吗？

因为我们对信息结构的相对尺度没有直观的概念，所以计算机成果与自然结果的比较，走过了一段艰难的时间。专业和普及的出版物都在

连篇宣扬人的计算能力很快就要赶上自然的复杂性了。例如，它们反复宣称计算机将最终认识人的情感和语言，计算机能在复杂的生命与我们只学会了编目的简单DNA序列之间搭起沟通的桥梁。

为了刻画我们在这个问题上的无知的本质，可以提出这样的问题：自然进化是一个笨拙、缓慢、低效的过程，还是某个像自然组装的超级计算机的结果？——在某些情况下，那个超级计算机甚至能在量子水平上运行，它能自我优化，从而在几乎最可能短的时间里产生具有不可约减的复杂性的结果。这两种情形是所有可能现实的两个极端。真实的情况我们还不知道，也许界于二者之间。我个人倾向于第二种：进化在完成不可约减的艰难使命时可能是非常高效的。[103] 不过，大多数当代关于科学与技术未来的对话却似乎接受了另一个极端——再经过三五十年摩尔定律的魔力，我们的计算机就超越自然了。

在线路和协议局限下的20世纪中叶的计算机科学支配着计算和生命系统的文化隐喻。例如，博尔赫斯（Jorge Luis Borges）描绘了一个幻想的图书馆，能囊括所有已经写了和将来可能写的图书。[104] 假如你有幸正好生活在一个能装得下它的宇宙（我们不是），那么你需要在星际飞船上耗尽无数代人的生命才能到达存放你需要的那本书的地方。学会以传统方式来写好一本新书要简单得多。道金斯也想象过类似的无限的所有可能动物的图书馆。他想象那只看不见的盲目进化的手在轻轻抚过那个图书馆，为每一个生态小环境寻找最合适的动物。[105] 两种情形，作者都沾染了20世纪那不充分的计算机科学的隐喻。新的计算机科学尚未建立，不过我们至少可以猜想一下它可能的样子。

新的计算机和信息科学将包容一个遗留[106] 系统的理论。复杂因果系统的空间构形太大了，不能理解为无限的图书馆，因为永远也不可能有足够的时间和精力来有效地浏览它们。例如，考夫曼喜欢宣扬我们的宇宙还不够老，不足以探索所有可能的（哪怕是非常小的）蛋白质。所以，复杂系统积聚着太多遗留的东西，限制了进一步的构形空间的找寻。我们必须学会丢掉我们能克服遗留的幻想。当老练的技术专家们提出给人类的新陈代谢或大脑结构添加基本元素的时候（确实有许多这样的建议），正是那些幻想在起作用。

还有一点值得我们思考的是："遗留"和"语义"是不是一样的东西？"语义"描述的是以协议为基础的系统的语法壁垒特征以外的任何神

秘事物。例如，人们总说自然语言系统是在发展中的，然而缺乏对语义的理解。遗留系统在信息系统中产生一个永恒不变的语言环境。遗留系统是复杂的。在减小系统的构形空间中，它们起着透镜的作用，增强了信息量的因势。

如果在大街上陌生人拦住你问有没有火柴，你说“有”，不会有什么麻烦；同样的话在婚礼上说后果就不同了——至少一般是这样。婚礼是祖宗留下的东西，积淀着不可能轻松抹去的历史。同样，DNA只有在胚胎环境下才可能有意义；我们前面说的思想实验中的聪明的异类，肯定不可能从一段孤立的DNA片段获得足够的信息来重新创造生命。

新的计算机科学也许会吸收一种粗粒化的方式，把自然系统作为信息系统来认识——这种方式超越了香农提出的细粒化的例子。人们常说，20世纪末，我们的物理学知识足以解释所有发生在生命系统里的孤立事件（例如化学键），所以我们现在才必须进一步去认识复杂系统。这话好说却难做。我们需要学会根据因势来解析自然系统。任何时候，特别是在生命系统的情形，只有很小部分的系统物质和能量能对那个系统的未来产生重大影响。即使那样，影响的程度也有区别：例如，与比皮肤表层或者身体其他任何部位的细胞的微小改变相比，同样微小的大脑皮层的改变有着更大的意义。

考夫曼提出，生命可以定义为既自我复制又做卡诺循环（把能量转化为功的经典模型）的过程。这至少指出了一条可能的分析自然系统的途径。每个卡诺循环都关联着某种类型的调节器，即系统中负责重新启动循环的部分。这些调节器比循环中的其他部分有着更大的因势，就是说，调节器的微小改变比同一系统其他东西的改变更可能导致系统的崩溃。这种分析方法能否用来将自然系统粗粒化地理解为信息系统，现在还不清楚，不过应该能发现某种方法。

如果我们发现了正规而且通用的粗粒化方法来把物理系统解析为因果的信息结构，我们可能获得一种同时包含了计算和能量部分的复杂性的测度。例如我们可以问，一个系统探测自身内部的因果结构需要多少费用，因为这种因果链的环节是可以物化的。我们有过简单自然系统的粗粒化解释的经验，它也许能帮助我们去模拟自然进化的遗留。在未来50年，假如幸运的话，我们大概不但能描写DNA是如何工作的，出现的是什么DNA（这正是我们刚刚开始做的），而且还能有办法描绘约束

DNA改变的中间层次的复杂性。总的说来，我们也许能在一定程度上学会以进化的观点而不是分子或生命的观点来认识我们的世界。

50年后，生物学和医学可能会变得有点儿像我们今天看到的地理学。大多数研究领域将被开拓出来，神秘的地方将越来越少。可惜的是，单画地图只能很有限地促进两地间的旅行。同样，能解释今天的生物学有哪些奥秘，并不等于我们就能自动地控制它们。相反，我们很可能发现哪些生物学方面的复杂性是不可能约减的。进化经历那么长的时间才达到一定的构形可能是有原因的，而我们也许会发现并不存在什么捷径。这可能就是信息科学前缘的一个遥远和最后的港湾。在从经济到农业的大量探索中，我们将被困在复杂性的顶点下——制造更大更快的计算机也未必能打破这道壁垒。我们开始意识到，复杂性的顶点是对我们能力的最真实的约束。我们现在还不知道它们在哪儿，不过50年后会知道的。

拉尼尔（**Jaron Lanier**）

拉尼尔（Jaron Lanier）是计算机科学家和音乐家，最有名的是他在虚拟现实中的工作。他是“国家远程沉浸行动”[107]（一个研究下一代互联网技术开发和应用的大学团体）的主要科学家。

流动的信息

D.格冷特

David Gelernter

未来50年的计算机技术会发生什么事情？半个世纪以后我们会在什么地方？今天，信息在线上流动；不久，“伟大的理性化”就要开始。跟其他所有的技术产业（如铁路、汽车、广播、电视）一样，信息产业也会树立新的标准形式。这些形式不会跟我们今天的商务软件的应用发生关系。重要的在于信息本身是如何排列的；问题不在网页浏览器的标准，而在网站的标准。（网本身会落伍，而基本的思想还在。）

信息的标准形态将表现为我所谓的信息束。信息束跟书一样重要。它不会取代书，而将为虚拟世界提供一个同样强健、坚固而简单的结构。它将重新塑造我们的文化生活。今天我们大约80%的工夫在关心形式（以不同的方式），20%在考虑内容。50年后，这个比例将颠倒过来。

最重要的信息是刚从网络传来的实时的信息——它告诉我们在哪个时刻、在什么地方正发生着什么事情。今天，“什么地方”指地球上的某个地方——如办公室、学校、参议院、市中心。不久以后，它将意味着虚拟空间的某个地方。连绵的无处不在的虚拟世界将取代今天断续的混沌的互联网络。例如，纽约的股票交易正在从地球的空间向着虚拟的空间转移；在未来的半个世纪里，几乎所有其他机构都会跟着转移进来。（这是我1991年在《镜像世界》中第一次提出的，现在仍然坚持。）不必从床上爬起来，不必离开舒适的椅子，不论在办公室、在学校、在商场还是在你喜欢的世界，你都能与正在进行着的生活发生联系。但你也一定会走出来看看。实际上，为了适应这种新的文化形势，自然和社会的景象都将焕然一新。

在勾画这些巨变之前，我们先来看看指引它们的一些自然法则。

1）在技术世界里，软件（而不是硬件）决定着技术的状态和变化的步伐。技术进步的速度不依赖于我们发明的网络（或计算的蛋白质等）。它依赖于我们设计的软件结构。假如你设计了一种新的编排信息的方法，一种新的软件结构，支持它的硬件也最终会开发出来；假如你开发了强力的新硬件，它本身是一点儿用也没有的。十秒钟以前，我的秘书拿着一本关于计算机未来的新书走进来，要我写几句夸奖的话。它跟绝大多数这类图书一样，也是谈硬件的。关于未来软件的书几乎没有。看来，没人知道软件的未来。

考虑未来半个世纪的技术会是什么样子，我们很容易想到飞快的网络、分子和光学的计算机、新的数据传输媒介以及其他的硬件奇迹。它们是重要的，令人惊奇的，然而它们本身却跟我们不相干。半个世纪以后的技术状态依赖于我们发明的软件。

看几个相关的例子：20世纪80年代中叶以来，计算机硬件在以惊人的速度飞速向前。可结果呢？在计算方面我们比1985年到底好了多少

呢？我们黄油加面包地同计算机打交道，跟16年前几乎是一样的。2001年的文字处理器并不比1985年的模型好；它耗尽千百倍的储存和计算能力，却没做任何跟以前不同的事情（不管怎么说，没做什么重要的事情）。我们的数据表基本是一样的，我们的电子邮件也是一样的——用的人多了，但邮件本身还跟1985年的一样。我们的台式电脑、文件系统、图形用户界面（如果有Mac’85）——一切都还是它们15年前的样子。计算机给我们的生活品质带来的唯一巨大变化是网络，但网络是软件而不是硬件产生的。

今天，软件停滞了；于是技术产业也停滞了。硬件革命不值钱了。为了摆脱停滞，我们需要软件的革命——我敢打赌，我们会有一个那样的革命。

2）替代定律：社会取代一样事物是因为发现了更好的事物，而不是更新的事物。不要指望50年后凡事都会不同。基本的东西还是一样的，这似乎是显然的，然而不是。

去年夏天，《纽约时报》有一个头版标题说“电子图书时代，预言为时尚早”。在那一年之前的2000年8月，巴尼斯-诺贝尔（Barnes&Nobel）、微软和其他几家公司正式宣告电子图书时代来到了。他们的伟大预言错了。图书延续了两千年，是因为它好，而不是因为计算机工程师没有考虑过取代它——我们应该能相信这一点吧？正如我1999年写的（也在《纽约时报》），“拿计算机来取代书，就像拿塑料花来取代插花。”我认为，书是近两千年来最伟大的设计。

不过，令人惊奇的是，《纽约时报》在去年夏天的那种论调，我们以前就听说过了。早在20世纪70年代，施乐公司就已经宣告了图书的死亡和“E-书”的兴起（大概是这个意思）。我们可以肯定一点，再过十年或者二十年，《纽约时报》大概会有另一个头版大标题：“任由专家评说，图书还会保留”。（知识分子就是这样擅长反复利用别人的错误。）

50年后，我们还会读印在纸上的书，看画在布上的画。如果运气好，你今天可以听贝多芬，看阿斯泰尔（Fred Astaire）的电影[108]，50年后，（假如你还活着）你仍然可以那样。

3）有形的总是战胜无形的。图书战胜电脑屏幕，并没有什么根本的理由，也不因为书是雅致的给人美感的东西（当然是的）。图书战胜屏幕是因为它很实用。它生来就方便携带，翻阅，浏览，可以在上面写写画画，比屏幕上的东西更好读。但在50年或更短的时间里，这个“有形战胜无形”的定律，将使我们今天熟悉的许多事物走向消亡。

买东西是一个典型的例子。大家都同意，友好的“中央大街”小商店比大的购物中心更好，而购物中心比批发商店好。没人喜欢批发商店，但是许多人还是去那儿。谁能在亚马逊网上书店享受到买书的乐趣呢？那些书不能翻，也摸不着。不过，也有的买书人不去书店而上亚马逊，是因为它很方便、选择多，有时还便宜——这些有形的好处在任何时候都能超越无形的东西。

大学如果不是太得意自满，将为这个定律感到困惑和恐惧。在线教育已经全方位地在眼前出现了。如果我们想学的所有课程都能在网上学习，而且在线课程软件的质量在逐年提高，那么大学还凭什么来证明它的存在价值呢？大学能出卖的是无形的东西。它们提供了无形的校园经历——让你直接面对你的老师，更重要的是面对你的同学，面对你的校园。于是，50年后，大约95%的大学将消失。最好的一些学校也许还能坚持下来，因为它们确实有值得夸耀的实在的东西——声誉，这能给人们带来好的工作和财富。不过它们当然也会发生改变。例如，英语系原是为向学生传播伟大的文学而设立的，现在成了奢侈的地方。很多英语系在今天还坚持认为没有像文学那样伟大的东西。过不了多久，人们也许不会说，“是那样吗？那么你也可以教学生一些老古董呀。”他们会说，“是那样吗？那么我们不再需要英语系了，不是吗？”

当然，小学校也将彻底失败。

4）“古尔德定律”：技术终究只是手段，而不是目的。这位伟大的钢琴家已经过世差不多20年了。古尔德喜欢技术，也掌握着技术。20世纪60年代初，他曾大胆预言录音将取代现场演出。他离开了舞台，走进了录音棚。[109] 他关心录音的每个方面——包括音乐的细节和技术的细节。随着20世纪杰出作品的流行，古尔德的录音一直是人们最喜爱的。狂热的古尔德相信技术能取代数百年的表演艺术传统。但他也拿最新的音频工程技术来为钢琴（有时也为管风琴和古钢琴）录音。

我们今天迷恋着技术（随便拿张报纸来看看吧）。这是一种不健康的迷恋；我们喋喋不休地谈技术，以逃避那些令我们感到紧张和愧疚的话题。如果谈技术，我们就不必谈艺术和科学，谈真与美，谈父母对孩子的道德和精神的（与金钱相对的）责任和义务。不谈道德和精神的平庸，我们可以谈财政和工程的辉煌。

技术是迷人的话题，也是深刻的话题，但（正如古尔德知道的）终究只是手段而不是目的。50年后，技术会比今天更普遍、更有威力——今天它也够威风了——但我们不会像今天这样依赖它。

未来几十年的某些硬件事实是显而易见的。计算机和大容量记忆芯片会变得很便宜，千百万的寻常人家都可以把它们安装在每一座竖起的大楼（不论私人的还是商务的）的框架里。你可以不时改换它们，就像换瓦片一样。

50年后（也许等不了那么久），互联网将被一个充满信息束的虚拟空间所取代。连通其中的一束，就等于连通一个物化的思想——你自己的或某个组织、某个机构的。你所依赖的信息——你自己的生活故事加上几百个你最喜欢的虚拟空间里同样的故事——将存在于你生活的任何地方，你走到哪儿，它们就跟到哪儿。你不需要带着这些信息结构，它们自己会在虚拟空间里穿行，就像你在海边漫步时跟着的一群友好的海豚。

信息束的结果之一是我们能获得极大的安全。每一个数据（当然，加了密的）将在整个虚拟空间复制几千次。每一个数据结构将分布——分散——在几千个分离的微电脑里。只有千百次独立而协同地闯入那么多电脑，才可能破坏或盗取数据。而你放在自己钱包里的私人密码钥匙，不但能解开数据的密码，还能将千百个分离的没有意义的数据片段结合起来形成一幅有意义的大图画——只有在你看它的时候，那图画才存在。

我们来做一个思想实验：想象一束光从一面墙壁的中央穿过，射到对面墙壁的中央。你站在屋子的中间，让光束正好从面前经过。当然，光只有在照到什么东西时才看得见，所以，一束光其实就是一束照亮的灰尘、雾滴或其他漂浮的微粒。信息束像钟那样“以时间的速度”运动：每一颗照亮的雾滴连续地从屋子的右端（将来）穿过中间（现在）然后

流向左端（走进过去）。在垂直于光束的运动方向上插进一块屏幕（就像网球拍），这样你就可以“收听”它了。这就是你的“光束调谐器”。光束以稳定的速率通过你的调谐器。通过观察屏幕，你可以在光束经过的时候看到它。

实在的信息束是一股流向过去的一条条信息组成的溪流。如果将C—SPAN的节目转化为信息束，我们可以把它想象成一帧帧配有“一点点”声音的定格图画。（在电视上看C—SPAN，它们是一幅紧跟着一幅连续打开的。）[110] 想象这些图画在信息束中展开，而你站在屋子的中间。右边一半的束是空白的；右边是未来，那里图画还没有播出。当C—SPAN播出一幅新的图像时，它就在屋子中间的“现在”显现出来，然后流向左边的过去。假如你的调谐器在“现在”或者无限接近“现在”的左边，你会看到一幅幅新图像的出现。于是，在这个思想实验里，你是通过在你的调谐器的屏幕上捕获一幅幅图像来观看C—SPAN节目的。

站在你左边的人也能接收同样的图像。他也看C—SPAN，但他离“现在”10分钟（比如说，在你左边五步的地方），他看的是过去了10分钟的C—SPAN，比实况晚了10分钟。C—SPAN向你的左边远远延伸，你可以在一小时、一年或十年以后看它，这要看你站在什么地方。那么你右边的未来的情形又如何呢？C—SPAN有未来的计划：它的节目时间表。它存留在信息束的未来部分。假如C—SPAN要在下周二上午10点播出一台西班牙语教学节目（Spanorama），它就在信息束的下周二上午10点那个位置做一个记号，那个记号平稳地流向“现在”。在“现在”，时间表不再存在，而是转化为节目一秒一秒地播放出来。

信息束之所以重要是因为我们可以把自己的“信息生命”储存在里面。我们还可以把一个机构储存在里面。

我们来看你的信息生命。这个束是你产生或接收的每个信息组成的序列。束的信息可以是一幅画、一段音响、一段录像、一份文件、一页传真、一个网页书签或者任何别的信息单位。（刚才讲的C—SPAN束的元素是均匀的；大多数真实的束却是非常不均匀的。）你的信息束的起点——在你远远的左边，而且还在平稳地远去——是你的电子文本的出生证。新的电子邮件正好在你的面前（在“现在”的线上）出现。你做过的所有文档都储存在束的左边、过去的某个地方，并继续流向远方。如果要在某个文档的新版本上做事，你可以把文档复制一次，放在“现

在”线上，然后在那里工作。（为了移动文档，改正你的Word处理器：把它固定在某个文档；文档移动到过去，不过还保留在屏幕上。）在你的信息束的未来储存着你的计划、约会和提示。它们都向着“现在”流动，经过“现在”线，然后流向过去，走进历史。

我还有好多没有讲；没有解释信息束的“为什么”和“怎么样”。不过关键的一点还在于信息束把握了你的生命的历史记录。假如你在“现在”线上审视你的信息束，那么每一点新的信息都将出现在你的眼前。（电话交谈也是它的一部分。）在信息束，你可以接近你的整个过去；你所能预见的未来也在它上面。（在众多的问题中，有一个尽管很小然而却很重要：你的医疗记录也在信息束上。他们是属于你的；你指定的任何人也可以在任何地方直接看到它们。）

未来的虚拟世界充满了信息束。

如果你在工作，你公司的“信息生命”也在信息束中流过。每个意见、每个通知、新的命令或者需要大家讨论的家庭问题，都会出现在“现在”线，然后流进过去。每个人都可以自己打开公司的信息束。只有你才能看到的你自己的文档和电子邮件，也散布在公司的文件中。讨论在信息束上进行，然后进入历史。规则在信息束上建立，命令也在信息束上发布。工厂的经理们监视着信息束，从信息束得到命令，然后付诸行动。公司的计划、会议、项目和最后期限等，都在信息束的未来。

信息束是公司的大脑。公司从来没有过“大脑”。公司的大脑不像你个人的，它是许多人而不是哪一个人建立的。它记录了过去、现在和未来，不会忘记任何事情。打开公司的信息束，你就走进了公司的大脑；实际上，你成了那个大脑的一部分。

中小学和大学也是这样的。一堂课或一门课程也像一个信息束，它把课程内容一点一点地传达出来。众多学生可以同时在线上用功，每个人在不同的点；而老师在一旁关注整个信息束，在需要的时候更新材料，听取他们提出的问题。一个电子校园是一个信息束。什么是“校园生活”呢？就是不断进行的讨论（加上物理接触，那需要你自己来做）。在校园的信息束上，同时回响着千百个讨论，一会儿分开，一会儿又汇合；有时候，学校本身也“发言”，发布一个个分离的信息，例如通知和命令。所以这些信息单元都散布在讨论中间。学校的计划和课程

表张贴在信息束的未来，校园的讨论则向后流进过去。校园不再是时间里的一个点，它的整个历史都在信息束上，你可以在任何地方（甚至在任何时候）进入它……校园（或公司）也不是空间的一个点。不论在什么地方——坐在起居室或者躺在海滩上——你都可以浏览信息束，走进一个共同的头脑。

市场是信息束，是买卖人集合的地方。

但是多数人都喜欢同其他人在一起。我们不想呆在家里。50年后，人们去某个地方、加入某个人群，都是因为他们愿意，而不是因为不得不去。学校将成为邻居孩子们的随意集合。每个孩子都能打开一个独立的信息束——坐在“邻里学校”教室里的20个小孩可能同时加入20所不同的学校，但他们可以一起午餐，一起在教室外面蹦蹦跳跳；任何有责任心的成年人（不论有没有教师资格）都可以照看他们。同样也可以出现“邻里办公室”。你可以在一间小办公室里跟几十个人一起工作，他们也许来自不同的公司，但是仍然愿意在一起度过工作日。

形成我们今天景观的办公大楼将不复存在，大大小小的商店即将退出舞台。（电子商务似乎暂时还悬着，但别忘了今天的网站是多么可笑和原始。例如，网上书店的图书为什么不让我们翻阅呢？这是没有理由的；为什么要我们在每次访问新网站时都熟悉新的界面设计呢？这也是没有理由的。不管怎么说——尽管结果也是好混杂的——商务和教育正在无情地走进虚拟空间。）在世界延伸的信息束的最终结果是使邻居像在19世纪那样重要。人们需要房子，需要方便普通的邻里聚会的公共场所。除了大博物馆、大主题公园和大购物中心，我们不再需要城市。更糟糕的是，城市是人类最伟大的艺术品。不过，当我们不再需要它们时，一定会更欣赏它们。

在未来50年，我们驾驭的东西会不会更少？不，我们能驾驭更多。明白地说，我们喜欢征服。我们越是富有，越能做更多我们喜欢的事情。

所以，半个世纪后的世界跟今天有不同样子、不同的工作方式。它将更加富裕；它将拥有更先进的技术。它也许还会比任何时候都更加幸福一点。

格冷特（**David Gelernter**）

格冷特（David Gelernter）是耶鲁大学计算机科学教授，镜像世界技术（纽黑文）首席科学家。他主要研究信息管理、平行规划（parallel programming ）和人工智能。在Nicholas Carriero and Gelernter’s Linda系统（1983）中引进的“数组空间（tuple space）”是世界许多计算机联络系统的基础。格冷特博士是以下几本书的作者：《镜像世界》（Mirror Worlds ）、《机器的缪斯》（The Muse in the Machine ）、《1939》《描绘生命》（Drawing Life ）和《机器美人》（Machine Beauty ）。

思想，大脑和自我

J.勒杜

Joseph Ledoux

年轻的弗洛伊德从研究神经系统开始他的科学生涯。他相信，对大脑功能的认识能揭开精神生活的秘密。他很快意识到已有的大脑研究工具不足以实现他的理想，于是转向了纯粹的心理学方法。在他以后的那些年里，神经科学已经茁壮成长起来了，它的发现一定能令弗洛伊德感到惊奇。不过仍然还有许多需要研究的东西，下面讲的是我们可以预料将在未来年月出现的进展。

读脑

我们的感觉、记忆和情感是如何通过大脑活动的？神经科学研究已经卓有成效地揭开了这些秘密。许多研究涉及人类以外的生命，特别是老鼠和猴子。虽然这种方法能让我们充分了解人类与其他动物所共有的大脑功能，却不能使我们认识人类大脑的独特性。研究有脑损伤的人有助于填补这个空缺。虽然脑损伤问题的研究关心大脑的正常功能，实际上也同样关心大脑如何补偿失去的功能。

新的技术使我们能够研究人类大脑的正常功能，而且有望在新的水平认识大脑与思维的关系。具体说，功能磁共振图像（fMRI）的出

现，为研究者深入人类大脑、观察它在产生心理活动或经历一定事件时的活动，提供了安全实用的途径。目前所做的多数图像研究都是为了发挥这个技术的作用，证明它跟更传统的方法揭示了相同的大脑功能图像。现有的许多发现也就这样与来自试验动物的大脑研究相互参照。如果对特定的大脑系统的行为没有基本的认识背景，图像发现的东西就将悬在认识的真空。举例来说，老鼠和其他哺乳动物的研究表明，在发觉和感应危险时，扁桃体（颞骨附近的小脑的一叶）是大脑网络的关键部分。在这一点认识的指引下，研究者后来发现，扁桃体损伤的患者对危险的识别很迟钝；当人们暴露在威胁的刺激下时，正如fMRI图像所反映的，扁桃体区域会活跃起来。在这样一些方面，动物的研究在前头铺平了道路。

重要的是，研究的物种要与提出的问题相匹配。例如，我们赖以把握信息和做事情的短时记忆，是人类思想的重要的基础过程。我们知道这个过程与人类所谓背外侧前额皮层的大脑区域有关。老鼠没有背外侧前额皮层，所以不适合做这类关于记忆的研究。猴子有背外侧前额皮层，而且，关于这个区域在短时记忆过程中的作用的许多东西，都是通过对猴子的研究而发现的。但是，人类思维的一个重要方面是言辞的短时记忆，这个功能不可能在人以外的其他物种进行直接研究。最近，fMRI研究对说明人类大脑如何实现言辞的短时记忆，发挥了重要作用。

将来用fMRI或其他方法——包括记录活动的方法、刺激选择大脑区域的方法和诱导活动的方法——来研究人类大脑，可能分化为三个大的领域。第一是最普通的：对已经有所认识的过程，如感觉、记忆、情感、语言和短时记忆的神经组织，我们将获得更多的认识。第二是更多地发现这些过程是如何在大脑中相互作用的。这些发现将使我们在更加广大的系统水平上认识大脑功能的概念，而且至少能使我们走近一个理论的开端——这个理论关心的是大脑如何产生思想，而不是特定的精神过程如何发生作用。这样的研究已经开始了，可是太少。

第三个领域也许是最重要的。几乎所有大脑功能的研究都集中于多数人的大脑在多数时间里的典型行为方式。为了得出一个标准，这些研究都涉及许多对象。一旦对这些标准的功能有了牢固的认识，我们就能追问个体差异如何决定表现自我和个性的那些独特性质。这些问题需要多少有些不同的研究方法——在特定的个体做大量的观测，而不是对众

多的对象做同一个观测。

在人的大脑和思想中进行着怎样的活动，现有的技术为我们提供了有力的评估工具。随着这些技术的进步，我们不得不面临这样一个问题：我们的社会是否能够承受研究可能带来的一切？假如我们能看透人的大脑，能看清一个人在想什么、感觉什么——从而知道他是不是想谋杀、猥亵或者强奸——知道了这些东西，我们该怎么办？

控制记忆

每当我们形成记忆的时候，实际上是在调节大脑的某种联络——神经键的联络。不论我们平常记住早晨穿的袜子的颜色，还是本能地记住母亲的声音，记忆都是调节神经元之间的联络的过程。简单说来，事情是这样的：在事情发生过程中活跃起来的神经元，会经过一定的化学变化，那些变化能激活基因，从而启动活跃细胞里的蛋白质合成。然后，蛋白质转移到活跃细胞的活跃的神经元突触，在那里改变它们（而且只改变它们）从与之相连的神经元接受信息的能力。记忆就在这样的变化中形成了。据我们已有的知识，我们预料在不远的将来可以通过不同的方式来控制记忆。

现在人类的寿命长了，许多人都面临着与老龄有关的记忆问题。这些问题最显著地表现在阿尔茨海默症和其他某些神经障碍患者的身上，但没患什么特别脑疾病的老人，记忆力也会衰退。眼下，科学家正在研究形形色色的动物的记忆，如海蛞蝓、苍蝇、老鼠、兔子和猴子，拿那些知识来为增强我们人类的记忆力开辟新的途径。例如，我们已经明白，许多形式的记忆都跟神经传递介质的谷氨酸盐及其受体有关。于是，提高记忆的方法之一，就是开发能促进谷氨酸盐传递的药物。而且，化学离子（特别是钙）从谷氨酸盐受体流向神经元，也是记忆形成的重要步骤。钙的增加，能激活分子，而激活的分子又能激活基因。这样，开发促进大脑细胞中的这些过程的药物——就是说，提高它们激活基因的能力，形成更多的蛋白质，从而稳定作为记忆基础的神经元网络——为我们提供了增强记忆力的另一条途径。

不过，对于像患阿尔茨海默症那样的神经疾病的人，该如何修复他们的大脑呢？最近发现，在成人大脑的海马状突起区（对有意识记忆起着核心作用的一个大脑区域）长出了新的神经元，这为我们带来了新的

希望。假如能够通过什么方式让这些细胞联络参与到这样的退化记忆回路中来，记忆功能也许还能复原。假如联邦政府能让研究者们放开手脚，更自由地做干细胞的研究，阿尔茨海默症之类的疾病可能根本不会在易感人群里出现。

脑科学可能产生重要影响的另一个领域，是阻止或者消除人们不喜欢的记忆，特别是那些创伤的记忆。那种记忆是形成创伤后压力障碍的主要条件。如果能尽快把它们忘却，压力可能在一定程度上得到缓解。研究者们已经提出了在记忆形成和稳固时改变记忆命运的方法，这有助于开发某些药物，在高度紧张的事件发生后立即服用这些药物，也许能预防创伤记忆。但是，因为记忆只需要几个小时——合成和利用蛋白质的时间——就牢固形成了，所以这个方法应用起来还有许多限制。不过，我们大概还有别的方法。

对老鼠的最新研究表明，如果大脑记忆区的蛋白质在回忆过程中被干扰了，那么已经形成的特定记忆也可能被破坏。但是为了能消除创伤记忆，同时完整地保留其他记忆，我们使用的药物必须针对创伤记忆所在的区域。这就需要我们找到在创伤后遗症中创伤记忆形成的位置，还要通过某种办法把药物限定在那个区域。我们将简单考虑这些问题。

当然，即使能淡化或者抹去人们头脑中的烦恼记忆，也不是轻而易举的事情。假设有那么一个大屠杀幸存者，几十年来生活在尸横遍野的记忆中。这些记忆显然已经成了他个性的一部分。尽管他可能为这些记忆感到痛苦，但是如果把生命里的这一重要插曲抹去了，他将成为怎样的一个人呢？

科学进步有时会成为日常生活的一个部分。于是，我们也许有那么一天，能拿非处方药来为特别的经历在头脑里留下特别强烈的印记。假如你想把生日晚会或婚礼生动地记下来，在晚会前，你可以服一粒能更有效激发谷氨酸或别的什么分子的药丸，那么，晚会上发生的事情就会一幕幕地刻进你的脑子。

这样重塑记忆听起来似乎很牵强，实际上我们一直在想办法增强经历的情感作用，让我们的回忆更生动、更久远。服药不过是为了同一个目标的另一种方法。在婚礼上，给新娘送一瓶药丸不如送一束鲜花那么浪漫，但是药丸也许能更有效地达成心愿（记住一个值得纪念的夜

晚）。当然，为了保险，你送了药还可以送花。

忘忧草

麦克白渴求“一剂甘甜的忘忧药”来根治痛苦[111]，我们今天真的有许多药能很好地用来医治抑郁或其他心理疾病。但药物的使用会产生一些副作用。50年后，或者更早一些，药物将针对大脑的病态网络发生作用，而不会影响其他部分。为了做出那样的药物，还需要几点进步。

首先我们需要更多地认识一定的病态究竟破坏了哪些大脑网络。在这方面，脑图像正开始发挥作用。研究正在向我们表明，抑郁者、焦虑者或精神分裂患者的大脑跟没有这些病痛的人有着怎样的不同。不过，为了认识这些区别的意义，我们还需要更多地了解那些病态区域的正常功能。

例如，根据现有动物和人类的数据，我们有理由假定，跟恐惧相关的疾病（急性焦虑症、创伤后遗症、广泛性焦虑症、惊吓、恐惧、妄想型精神分裂），源于大脑恐惧网络的正常功能和它跟其他网络相互作用的改变。我们曾经讲过，扁桃体是这些网络的关键部分，扁桃体功能的改变也许能解释焦虑症的某些方面。具体说来，当扁桃体对危险过分敏感，能从别人可能忽略的境况觉察出危险并产生保护性反应；或者当扁桃体活动过分强烈，对同样威胁能产生比别人强烈得多的反应，都可能出现过分的不适当的恐惧。这两种情形可能来自遗传，也可能来自创伤或其他紧张经历，也可能是两种因素的某种综合。而且，不论哪种效应，都能通过与扁桃体相连的其他大脑区域对扁桃体功能的调节来加强。不同的情形，可以用扁桃体内部或扁桃体与大脑其他区域之间的记忆回路的不同改变来解释。如果脑图像研究证明了扁桃体（或其他大脑区域）确实在焦虑症的情形下遭到破坏，那么，弄清本区域的功能和它与其他系统的相互作用，对开辟新的治疗方法将有着根本的意义。不过，即使目前的脑图像研究说明了人类的某些病态（如焦虑）牵涉着一定的脑区域，动物研究对我们在细胞和神经元的水平认识那个区域的详细的神经作用机理，仍然是很重要的。新的更好的药物的开发，最终也依赖于那个水平的知识。

一旦人类图像研究证明了特别的病态牵涉到一定的大脑神经网络，而动物研究详尽揭示了那些网络的组织，我们就可以期待专门针对那些

受伤害网络的药物。方法之一是利用分子遗传学的进步：假如我们能识别某个只在表现于扁桃体或者只能以某种特别方式在那儿表现的分子，那么我们就可能拿这样的分子做钥匙去打开新药的大门。就是说，药还是口服的，仍然可以在血液里流向大脑的多个区域；不过，因为药物分子的包裹，它们在多数区域是不活动的。只有在遇到分子“钥匙”时（在这个假想的例子中，只有扁桃体才有那样的分子），药物才能活动。这样的药物能帮助扁桃体修正反常行为，而不会影响大脑的其他区域，从而也减少了药物扩散产生的讨厌的心理学副作用。可是，因为扁桃体也参与了“正常的”大脑功能，所以我们面前的真正挑战是寻求某种能有选择地对病态功能进行攻击的途径。

扁桃体辩护

扁桃体跟许多大脑区域一样，在我们不知不觉中发挥作用。我们可以知道扁桃体的作用结果，却不可能有意识地走进它的内部。因为扁桃体可以被激发起来表现出无意识控制下的情绪反应，所以它很有可能会无意识地犯罪——做出清醒的人永远不能宽恕的事情。

这种可能性没有逃过法律。法律体系中早就认识到了“情绪犯罪”，即一个本来守法而且有理智的人在失去理性和心智时所犯下的罪行。“扁桃体辩护”为这类问题找到了神经学的理由。随着我们对大脑活动越来越多的认识，律师对我们的发现有了更多的了解，基于神经学的辩护会越来越普遍。所以，还是让我们具体来看看扁桃体辩护是怎么回事。

首先，扁桃体辩护不能同另一个相关的问题——我们可以称它“病脑辩护”——混淆起来。在病脑辩护中，问题在于一个人犯罪是因为他的大脑发生了生理病变。相反，扁桃体辩护的基本思想是，扁桃体是在无意识状态下正常地控制着情绪行为。于是，一个罪行可能是扁桃体在没有意识思维时犯下的。扁桃体当然也可能控制与一定刺激条件下的意识控制无关的攻击行为。然而，为了能实施扁桃体辩护，还需要满足几个原则。

扁桃体的一个重要功能是在面临突然的危险时迅速产生自我保护反应。但是假如刺激经历了一定的时间，而且能被意识所察觉，那么行为将更多地受通过大脑皮层的更高级的思想过程的控制。另外，扁桃体控

制下的各类反应是以一种固定的方式产生的本能的（“硬联接的”）、简单的、迅速的反应——就是说，在同一物种的所有成员中，这些行为都是相同的。所以，如果某个行为是故意的、表现相对迟缓的（经过几秒而不是几毫秒），有着一系列复杂的动作，而且不同的人有不同的做法，那么这样的行为很可能不是扁桃体直接控制的。扁桃体可以间接影响或协调这些更复杂的反应，但它们终究是其他大脑系统的事情。这些事实说明，能借扁桃体成功辩护的罪行，必然是相对单纯的、本能的、一成不变的反应行为，它几乎是在刺激发生的同时产生的，没有任何的预谋。

我怀疑没有几个案子能满足这些准则而赢得扁桃体辩护。不过，我们越来越清楚地发现，扁桃体之外的许多大脑系统在无意识地发挥作用——甚至意识本身也是大脑网络无意识行为的结果。这就产生一种可能：扁桃体辩护虽然在名义上错了，但在本质上却仍然是成立的。然而，我们是否需要重新考虑人类责任的本质和局限，还有赖于未来关于大脑有意识和无意识控制之间的平衡的发现。这些发现可能也在未来50年到来。

勒杜（**Joseph Ledoux**）

勒杜（Joseph Ledoux）是纽约大学自然科学中心Henry和Lucy Moses科学教授，长期以来寻求从大脑生物学状态的角度来认识人类情感。他的研究强调学习和记忆在情感经历中的作用（而不是遗传的先决作用），寻求将情感经历的记忆与神经元的连接联系起来。他的最新著作是《突触的自我：大脑如何适应我们自己》（Synaptic Self: How Our Brain Becomes Who We Are ）。他还写过《情感大脑：情感生活的神秘基础》（The Emotional Brain: The Mysterious UnderpinningsofEmotionalLife ），与Michael Gazzaniga合写了《完整的头脑》（The Integrated Mind），与W.Hirst编辑了《精神与脑：认知神经科学对话》（MindandBrain: DialoguesinCognitive Neuroscience ）。

我们何以如此：来自**2050**年的观点

J. R.哈里斯

Judith Rich Harris

作为儿童成长研究会健在的最老的会员（2月我就112岁了），我应邀来报告我们领域在过去50年——21世纪前50年——的科学进展。不过，在谈儿童之前，我想说几句有关我们这些老人的话。我们学会的人都知道，衰老也是一种成长。本世纪最重要的科学进步之一，在我看来，就是开发了能预防、甚至一定程度上能逆转阿尔茨海默症引发的大脑变化的药物。虽然我不想装着记忆力很好的样子——如果我忘了说什么你们觉得应该在这个报告里说的事情，我请你们原谅——但我今天能被请到这里来向你们报告，正好证明了那些药物的功效。

好了，现在我来谈指定给我的题目：儿童成长领域在过去50年有哪些科学进展。21世纪开始的时候，发展派学者们已经认识了很多对所有正常儿童都基本相同的那些成长环节。在认识儿童如何学会思想、说话、阅读等方面我们取得了好的成绩，至于是什么使他们成长为不同的人——为什么有的孩子善良、懂事，有的孩子好斗、冲动——我们却知之甚少。20世纪的发展派学者（那时他们称自己为“发展心理学家”）认为他们懂得了个体行为和性格差异的根源，但我们现在知道，他们多数都错了。所以，21世纪已经取得的最重要的进步，是我们认识了人为什么会不同，并且学会了利用那些不同。

讲这些进步之前，我想应该先来看看为什么这个领域在上个世纪里几乎没有进步。主要的原因是我们轻视了遗传学而运用了过时的研究手段。到2000年的时候，发展学家们才勉强承认婴儿并不都是天生一样的——每个人生来都有各自的特点，主要来自遗传。但他们仍然用着20世纪50年代出现的研究方法，而那些方法的基础却假定婴儿都有完全相同的起点！

你们看到了，20世纪50年代时，多数发展派学者的确相信新生儿都是一样的，于是认为他们长大以后出现的任何差异一定源于他们出生以后的不同经历，也就是不同的成长环境。在那样的假定下，在那个时候产生那样的研究方法是有意义的；但不幸的是在那个假定被抛弃很久以后，那方法还在继续流行着。

那个方法很简单。首先，所谓“社会化研究”领域的发展派学者会度

量环境的某个方面与儿童成长的某个方面。然后，他们寻找环境度量与成长度量之间的联系。接下来，他们报告他们的发现：例如，听父母读书多的孩子长大以后可能更喜欢读书；经常挨父母打的孩子长大以后可能更喜欢争斗；常跟父母说知心话的未成年人不容易陷入少年的各种烦恼。最后，他们根据这些发现，向孩子的父母提出忠告：如果想让孩子在学校学习好，就多给他读一些书；如果不想让孩子好斗，就不要打他；如果不想让孩子陷入烦恼，就经常同他说一些知心话。美国政府实际上花了很多钱让研究者们做这种研究，然后发表那些的忠告！

是的，我们今天可以笑话它，不过那时它却是很严肃的事情。用当时流行的词来说，20世纪末的发展派学者“活得很累”。[112] 他们没有面对这样的事实：假如基因对他们度量的结果有任何重要的影响，那么他们的研究成果就不可能得到解释。他们虽然承认儿童有基因，却不承认孩子继承了父母的基因，而且只是因为遗传他们才会在许多地方（如智能、胆量和为人）像他的生身父母。

为什么发展派学者经过了那么长的时间才承认这个明显的事实呢？毕竟，运用更好的方法所做的研究，已经产生了足够的数据证明他们正在走向错误的结果。那时候，所谓行为遗传学领域（现在我们更熟悉的是它的一些分支学科）的研究者们已经证明，20世纪发展派学者津津乐道的那些关系，几乎都可以在家族成员的遗传相似基础上得到解释。如果我们考察收养子女的家庭，那些关系也就不存在了。但是这些结果和理解它们的人所写的忠告，差不多都被遗忘了。

我就是那些人当中的一个，世纪之交的时候，我也为那些忠告白费了好多口舌。有人预言，我1998年的那本《教育假设》将成为“心理学史上的一个转折点”，很遗憾它成不了。[113] 那样一艘大船不可能一下子转过弯来。它正全速前进着，掌舵的是当时德高望重的大学者们，他们对现状感到非常满意。为了将它调转到新的方向，还需要大家来推动。如果我没记错，第一个推动是一本书：《家庭影响的极限》，作者是洛威（David Rowe），比我那本早几年。世纪之交不久，出现了宾克尔（Steven Pinker）的《空白的石板》[114]。几年后又有普洛明（Robert Plomin）的一本。再后来，还有伍兹（Eldrick Woods）和阿比盖尔·沃克（Abigail Valk）的。（说句题外话，也许你们有人不知道，伍兹以前还有过高尔夫球冠军的经历。至于阿比盖尔，当然了，她不仅做过学会的主席，也是我的孙女。）

不过最大的推动来自发展心理学领域的外面——抱歉，我这里说的是发展的科学。人类基因组的解密给遗传学研究带来了巨大的推动力，它首先使人们发觉，然后又使人认识了，微小的基因差异如何能在人的个性和认知能力上产生显著的不同。研究者最后面对着这样的事实：除非知道环境给儿童带来了什么特征和倾向，否则他们不可能明白环境是怎么影响儿童成长的。离开了对基因的控制，只关心成长结果的研究不能告诉我们任何事情。

现在，在寻找各种基因组合中通过扫描个人的基因组，我们可以直接控制许多类型的基因作用。不过，从长远来看，控制基因作用还需要运用更艰辛的方法，例如，研究收养的儿童或者双胞胎。尽管这些方法显然能产生更多的结果，老方法还是一直用到了2016年。就在那一年，美国政府终于下了决心，拒绝资助没有包含基因控制的任何关于成长的研究。这个决定革新了这个领域——不仅因为它结束了无用的研究，还因为好多老一辈的发展派学者在那个时候决定退休了。

当然，帮助把我们的领域转向新方向的，还有其他的因素。我要说的一个，是从古人类学发现得来的知识。最重要的一个发现是在2021年，在一个融化的冰川……哦，那是在哪儿呢？——大概在斯堪的那维亚的什么地方。在冰川里发现的是一具早期欧洲人的遗体，大约死了27000年。不过，引起轰动的不在于他是欧洲人，而在于他穿戴的东西。他的衣服是精美的厚毛皮做的，起初没人能认出来。好，你们大概知道，那原来是一个尼安德特人（Neanderthal）的毛皮——准确说，是三个尼安德特人的。这个发现极大改变了我们关于原始人类的演化和历史的观点。它说明了古人类学家早该认识的一件事情：尼安德特人是有毛皮的人。如果没有厚厚的毛皮，他们不可能在冰川时代的欧洲活那么久；随便在肩上披一张鹿皮是不可能抵御那种恶劣气候的；他们还没有发明针线，当然也不会把它缝起来。起初，人们很难接受这样的观点：我们祖先不仅把尼安德特人看成是食物的来源（那似乎是可以原谅的，因为尼安德特人也是那样看我们的），还是衣服的来源。[115]

尽管那个发现只提出了我们已经知道的事情——我们是掠夺者，是世界上最可怕的动物——最终它还是让我们更现实地来看人类的本性。那种浪漫的“高贵的原始人”的观点终于被扔进了垃圾堆。我们的祖先是原始的，是的，但一点儿也不高贵。如果做好人，我们也走不到今天了。

不过，我们显然还是可以在适当的条件下做好人。研究者们在确定那些条件方面已经取得了很好的进展，但是也还有好多事情要做。

关于儿童成长，我们知道些什么？

很抱歉，大家到这儿来，多半是对儿童成长环境的影响感兴趣，而我却讲了那么多遗传学的东西。不过，正如我说的，为了弄清环境的影响，我们先得剔除基因的影响。有了21世纪的技术和方法，研究者们现在可以做得相当精确。

这个研究说明，尽管环境对儿童成长有着重要的影响，但并不像20世纪的发展派学者想的那个样子。现在看来，他们的理论所依赖的大多数关系，要么是基因的直接结果——如父母和他们的亲生子女有着相似的基因组形式——要么取决于父母对孩子的行为的态度。有些十几岁的孩子对父母说的每样事情都瞧不上眼，他们的话一句也听不进，那么父母是不大会同他们说知心话的。如果真的这样，那些孩子是最可能陷入烦恼的。父母没能很好与孩子交心，孩子容易做出蠢事来，两者之所以有联系，是因为它们有着同一个根源：孩子的个性。

于是问题在于，我们如何解释孩子的个性？为什么有的懂事，有的冲动，有的讨人喜欢，有的令人讨厌？21世纪开始的几十年前，我们知道那不全是因为基因；不同个体的基因的差别大约只能解释个性差别的一半。但是基因之外的对个性发展的影响，我们几乎一无所知，因为，现在发现，过去多数的研究时间和金钱都浪费在了不可能有结果的问题上面。

当发展派学者终于发现个性与环境的相互作用时，事情才有了一点突破。人们很早就认识了环境强烈影响着个性——不同环境的人有不同的行为习惯——但是人从一个环境走到另一个环境，行为是有一定惯性的。20世纪的发展派学者错误解释了这种行为惯性。例如，他们发现有些孩子不论在家还是在学校都很烦恼，就匆忙做出结论说，孩子在学校烦恼是因为某些家里的事情。但是，一旦我们有了方法把遗传和环境对行为的影响区别开来，就可以清楚地看到，孩子在不同环境下表现的任何相似行为的倾向，几乎都是因为遗传对行为的影响。环境影响不会从一种情况转移到另一种情况（当然，如果两种情况相似，环境影响也会是相似的）。

这一点认识为我们解答了一个疑问：为什么儿童时代的家庭环境对儿童表现的行为方式似乎没有什么影响？在家里发生的事情当然是重要的，不过，为了认识它的影响，我们还得看看人们跟父母兄弟姐妹在一起时是怎么做事的。人长大以后，我们很少在那种情形下评说他的个性，自然也就不会表现出儿童时代的家庭影响。父母对子女的个性没有长久的影响，理由很简单：成年的人没有生活在父母的家里。

于是，为了弄清儿童时代的环境如何影响成人的个性，研究者们还得关心家庭外面发生在儿童身上的事情。事实证明，儿童在家庭外面遇到的几乎每一件事情——他们在学校和邻里的经历，老师和同学对他们的态度——都会产生影响。我们早就知道文化很重要，但我们也发现，文化只有在父母以外的其他人传播的时候，才可能产生长久的影响。如果完全靠父母来传播文化，孩子会觉得那只是他们家或家族的老习惯，在别的地方就不对了，或者没用了。父母发现，如果想向孩子灌输某种特殊的文化，他们只能把孩子放在那种文化氛围中，不但在家里，还要在外面。反正好多父母都在那么做，至于为什么有效，倒是一个有意思的问题。

我们还有很多东西要学。在世纪之交，我曾乐观地认为，到2050年的时候，我们能发现个性里多数非遗传变化的根源（那时我们把这些变化的原因归结为所谓的“非共享环境”）。但是直到现在，我们大概才勉强解释了其中的一半，还有大约四分之一的个性改变没有得到解释。我们知道，有些尚未解释的改变来自环境——来自发生在孩子身上的一些小事情，而那些事情的发生，我们很难研究，当然更难预料。但有些改变可以拿生物学来解释——尽管不是遗传的，却是生物的。即使基因完全相同的小宝宝（如孪生或者克隆），出生的时候也不会是完全一样的；他们的指纹会有细微的差别，大脑当然也是。哪些分子过程在大脑形成中产生了那些微妙差异，那些差异对个性会产生什么影响，是研究者们眼下正在研究的事情，不过工作才刚刚起步。

所以，我们还不可能对行为或个性做出什么有确定意义的预言。当我们需要预言别人行为的时候，一定会讨厌不确定；不过我们会同样为此而欢喜，因为别人也不可能预言我们的行为！

运用我们的知识

我们现在知道，发生在家里的事情影响孩子在家里的行为，发生在外面的事情影响他们在外面的行为。如果孩子的行为在家里出了问题，我们可以给父母讲一些更有效的培养孩子的方法，帮助他们。如果孩子的行为在学校出了问题，那就是学校的责任，我们也能帮助他们。例如，我们已经知道应该怎样保护那些弱小的、或者不太讨人喜欢的孩子，不让他们受恶人的欺负。

我们知道，孩子需要稳定的外在环境，经常搬家——不停地从这家邻居迁到另一家邻居，从一所学校转到另一所学校，对他们可不是好事。孩子特别需要经常在一起玩儿的伙伴，要不，他得老让新伙伴接受他，老让自己适应新伙伴的行为、打扮和说话。我们知道，孩子有多少父母是无关紧要的（性别更无所谓），只要父母的重新组合不破坏孩子在家庭外面的生活。20世纪中叶的人反对那些因为不愿继续生活在一起就分裂家庭的父母，而我们这些21世纪中叶的人，也不赞成那些为自己方便而带着孩子四处迁移的父母。

幸运的是，现在人们一般都在家庭能稳定下来以后才要孩子。由于生殖医学的进步，意外怀孕的事情几乎完全没有了。我们控制生育的能力在下降，问题在于我们没有更多的小孩。尽管多数政府在想方设法鼓励生育，但几乎世界每个角落的人口数量都在减少。

从孩子的立场看，这是大有好处的。当教师职位需要竞争时，意味着我们所有的老师都有很高的素质（也有很高的薪水）。孩子们在更小的学校和更小的教室里学习，事实证明它的好处远远超出了大家显而易见的教育的好处。过去，十几岁的少年往往在很大的中学读书，结果，他们可能分化为相互对立的群体：赞成教育的与反对教育的；爱运动的与不运动的；棕皮肤的与红皮肤的。这些影响一般是非常有害的。如果学校和班级都很小，就不大会发生那些事情。如果真的发生了，我们也知道该怎么做。

我们得到的知识也给父母带来了很大的好处。在20世纪的后50年，抚养孩子比以往和以后的任何时候都更艰难，因为“专家”让父母们觉得他们孩子的精神太脆弱，一步走错，就可能带来永久的伤害。父母们害怕使用自己的权威。他们很少体罚孩子；而是叫“暂停”，那样，孩子觉得难过，而父母更觉得难做。孩子的周围都是拥抱、亲吻、礼物、赞美和爱的承诺。孩子喜欢永远有人关心和赞美，但这个事实常被人误会，

以为孩子需要永远有人关心和赞美。在那些日子，人们看重所谓“自然”，但他们培养孩子的方式却要他们违背自己的自然倾向，装出一副爱的样子，其实他们通常并没有感觉到爱。他们不得不把自己的需要放到一边——连睡觉也顾不上。

孩子的培养方式，从严厉到溺爱，然后又回到严厉，经历着周期性的转变。我活了那么长的时间，看着它在两种方式间摇摆。50年前人们教育孩子的方式在我们今天看来有点儿可笑。今天的孩子没有得到过去那么多的口头赞美和物质奖励，但他们得到的却是纯真。在外面，日复一日、年复一年，孩子都跟同一群小伙伴们在一起。

令人惊奇的是，今天像这样长大的儿童，在许多方面竟然像远古部落或游牧社会的儿童，而不太像上个世纪末的典型美国家庭的孩子。我不能证明这有什么因果联系，不过，21世纪初空前多的儿童抑郁症，在今天明显减少了。

就借这句令人鼓舞的话来结束我的报告吧，谢谢大家。

哈里斯（**Judith Rich Harris**）

哈里斯（Judith Rich Harris）是作家，也是发展心理学家。她写过儿童成长的教科书，发现她过去告诉别人的许多东西都是错误的。她不再写教科书了，写了一篇文章宣扬她的新的儿童发展理论。文章发表在《心理学评论》（Psychological Review ），获1998年美国心理学会George A.Miller奖。哈里斯的书《教育假设：为什么儿童表现出那样的行为》（The Nurture Assumption: Why Children Turn Out the Way They Do ）曾入选1999年度普利策奖。

药、DNA和躺椅

S.巴龙德斯

Samuel Barondes

1950年，在一家法国制药公司Rhone—Poulenc[116]，一个化学家改变了一种治疗过敏反应的药物（抗组胺药）的结构，意外产生了一种能消除精神分裂症患者的错乱思想的药物。在几年的时间里，这种新药就以氯丙嗪（Thorazine）的名字闻名世界了，那是第一种真正的有效治疗精神疾病的药物。因为疗效显著，氯丙嗪为20世纪后50年开辟了一条精神病治疗法的新路。

氯丙嗪的巨大成功激发了其他制药公司的强有力竞争。20世纪50年代，在寻求更多抗精神病药物的过程中，人们偶然发现了两类其他精神病药。瑞士的一家制药公司（Geigy）生产了一种改进的抗组胺药，虽然对精神病不起作用，但后来证明是对付严重抑郁症的好办法。这种叫丙咪嗪（Tolfranil）的药，为现代许多抗抑郁药物的出现铺平了道路。接着，另一家瑞士公司（Hoffman—La Roche）生产了一种安定药利眠宁（Librium），对精神病也不起作用，但能减轻焦虑。紧跟着出现了另一种安定药重氮异胺（Valium），从60年代中叶起，在美国畅销了大约10年。

更令人欣喜的是，人们忽然发现那些药物影响着神经传递介质——在神经细胞间传递信息的一类脑化学物质。到20世纪70年代时，人们发现氯丙嗪阻断了一种叫多巴胺的神经传递介质的某些作用；丙咪嗪助长了其他几种传递介质的作用，包括去肾上腺素和5-羟色胺；重氮异胺则放大了另外一种叫做γ-氨基丁酸（GABA）的传递介质的作用。在每种情形，结果都是控制行为情绪的大脑信号发生了改变。

这些发现激发人们去寻找对神经传递介质有同样影响却没有那么多讨厌副作用的其他化学物质。结果发现了患者欢迎的一系列新药。其中最出名的氟西汀（Prozac）起初被认做一种延长5-羟色胺神经传递作用的化学物质，后来证明是治疗重度和中度抑郁症的有效药物。它通过抑制神经对其释放的5-羟色胺的再摄取，延长了5-羟色胺的作用（再摄取本来是中断5-羟色胺信号传递的正常方式），所以叫SSRI（选择性5-羟色胺再摄取抑制剂）。接着，与SSRI相关的药物也出现了，如盐酸舍曲林（Zoloft）、帕罗西汀（Paxil）、氟伏草胺（Luvox）、西酞普兰（Celexa）等。

随着对SSRI经验的增多，精神病学家发现这些药物还能帮助那些没有抑郁的人。如今，SSRI已经成为缓解无缘无故的恐惧（恐惧症）和无

法控制的焦虑（一般性焦虑症）的可靠途径——通过与安慰剂的对照试验相比较，证实了它有很好的疗效。

这样那样的新生药物的功效改变了精神病的治疗方法。在这些药物出现之前，多数精神病学家从纯粹的心理学角度来看待他们的病人，喜欢给他们做心理治疗。如今，他们的兴趣转向了大脑，而精神病治疗方法里常常至少包含一种药物治疗。几千万美国人在服用那些药物。

不过，尽管那些药物取代了氯丙嗪、丙咪嗪和利眠宁的作用，但也只不过是原来药物的改进，没有一种在根本上提高了疗效，而且还都有着令人讨厌的副作用。虽然我们已经广泛认识了它们对神经传递作用的影响，但新药的研究还有赖于20世纪50年代常用的那种试错法。

把今天的药物和心理疗法继续改进下去，恐怕也不可能迈出精神病治疗的下一步。未来的进步还有赖于我们去发现人类基因的改变和它们对大脑的影响方式。如果说20世纪前50年引导精神病治疗的是精神分析学家躺椅上的动听故事，20世纪后50年引导精神病治疗的是化学实验室里气味刺鼻的产品，那么，未来50年引导精神病治疗的将是我们对个体的基因差异的认识。

个体的基因差异的认识之所以能给精神病的治疗方法带来那么大的希望，是因为它能帮助我们回答一个基本的问题：什么决定着个体对痛苦行为的易感性？一个人过去的经历显然起着至关重要的作用。可是，为什么有的人能屡经精神磨难，而有的人容易沉浸在痛苦的状态呢？为什么屈服的人，有的堕入抑郁，有的无限焦虑，有的出现依赖药物的戒断综合症，生活在精神分裂的幻觉中？

我们隐约看到的线索是，所有这些痛苦行为模式都在家庭里世代相传。例如，我们来看一个人患精神分裂症的风险。大多数人表现出那种典型症状的概率是百分之一。不过，假如谁的父母或兄弟姐妹是精神分裂患者，那么他在一生中患病的概率就要大八倍。另一种精神病——躁狂抑郁症（也叫双相情感障碍）——的主要起因也是这样的。一般人患病的风险是百分之一，如果父母兄弟姐妹中有人患过这种病，那风险就会大八倍。抑郁和焦虑也是家族性的疾病。

不久以前，这些家族研究引发了一场大争论，争论的一方认为那是已知的家族反常行为模式的证据，另一方则把它们作为精神障碍倾向的

遗传证据。现在，多数认为，环境与遗传都起着部分作用。他们也同意，在评估遗传作用的下一步，最好是能发现有关基因的不同可能形式。

认识能达成一致，主要原因是我们发展了强有力的技术，能直接检验人类基因的不同可能形式，即所谓的等位基因或基因变量。这些从DNA结构的随机变化中出现的变量，决定着人类丰富的多样性，包括对某些特殊疾病的不同易感性。但就在不久之前，我们还只能推测它们的存在。新技术使我们能识别那些对人类品性有所影响的基因变量。现在，我们不再争论环境和遗传哪个更重要，而是集中精力去寻找哪些基因变量容易使个体患某个特别的疾病。

寻找的方法之一，是把家族中有那种病的人的DNA拿来跟没病的人做比较。假如只有患病的人才有某个特殊的基因变量，那么基因与疾病的相关性可能就是有意义的。假如同样的变量也只是出现在家族中许多其他染病的成员中，那么这个事实就更有力了。在某种情形下，这种可能性是很大的，从而那个变量的作用也就确立起来了。20世纪90年代，我们完成了人类基因结构的草图，影响对特定疾病的易感性的某些基因变量，就是通过这种方法识别的。有名的例子是三个不同基因的变量，每一个都会增大人们在50岁之前患某种罕见的阿尔茨海默症的风险。在一组家庭，罪魁祸首是一种叫APP的基因的变量；在另一组家庭是PS1；在第三组家庭则是PS2。[117]

那些增大患罕见阿尔茨海默症风险的基因变量的发现，激发了人们对精神分裂、抑郁、躁狂抑郁和其他精神疾病的研究。这些研究的巨大吸引力在于，它们不需要去猜想可能涉及了哪些基因，因为它们可以发现疾病与人类基因变量之间的关系。尽管过去的很多研究确实把焦点集中在具体的基因，特别是那些影响神经传递作用的基因，但我们对精神过程的基因控制还是知道得太少，如果说还有别的基因在起作用我们也不会感到惊奇。不幸的是，经过多年的努力，我们还是没有发现哪个基因确定地增大了某个精神疾病的风险；在其他常见疾病（如糖尿病和高血压）的基因研究也没有多少成绩。没有进步的原因之一是，决定那些疾病易感性的是多个而不是单个基因的变量的联合作用。虽然当前的技术使我们能相对简单地识别确实对某些风险起着主要作用的单个基因的罕见变量——如APP、PS1或PS2——但我们仍然很难发现那些只有在与大量其他基因共同遗传下来才能增大患病风险的基因变量。

随着我们对人类基因组认识的不断增长，那个困难很快就能克服。最近发表的人类DNA结构图迈出了关键的第一步。目前我们正在检验来自许多人的DNA样本，目的是对近3万个人类基因的共同变量进行识别和编目。这样，寻找那些可能共同对精神疾病的易感性发生影响的多个基因变量就简单多了。随着详细检验个体DNA的有效新技术的进步，那些寻找正变得更加简单。这些技术正在不断前进中，令人想起正在前进中的计算机芯片。同样，我们用以分析来自这些DNA研究的大量信息的计算方法，也在前进着。

收集和评估大量DNA数据的技术进步了，我们很快就能大规模地去寻找对特殊精神疾病的易感性产生影响的基因变量。随着DNA分析费用的不断下降，我们的研究可以走出家庭的小圈子，对成千上万没有亲缘关系的某种疾病患者的DNA样本进行分析。这样的考察应该可以识别相关的基因变量，而在个别患者身上，只能发现其中的某一个。

为了恰当地使用那么多的基因变量数据，不但需要联系病态的行为模式，还需要联系大脑的性质。许多新的技术，如功能磁共振图像技术，正开始用来评估个体大脑特定区域的功能。基因变量的模式与这样那样的研究结果相结合，能使我们分辨出更多的亚类病态模式，它们在今天的诊断标准里还没有分开，例如精神分裂或者抑郁症。

遗传信息与功能研究相结合，还能为我们指出真正的新药靶向。我们正沿着这条路线去寻求阿尔茨海默症的新疗法。目前治疗阿尔茨海默症的主要药物是通过延续一种叫乙酰胆碱的神经传递介质的作用来改善大脑功能，这种机制跟今天的其他一些药物（如SSRI）的作用是一样的。在阿尔茨海默症的罕见病例中分辨出APP、PS1和PS2，使我们能把注意力集中于其他可能的药物靶向。大脑中的这些被称作“分泌物”（secretase）的酶，参与产生了一种叫β-淀粉样蛋白的毒性蛋白质片段，这种蛋白的聚集也受着基因变量的不同方式的影响。几家医药公司正在研究抑制分泌物的新药，希望这些新药能减少β-淀粉样蛋白的聚集，从而阻止大脑的退化。

除了发现新的药物靶向，DNA研究还能认定某些特别的基因变量，它们能分别哪些人能从现有的哪些药物（如SSRI）得到好处，而哪些人不能。这些区别，可能是使个体有某种精神疾病倾向的特殊变量产生的，也可能是决定药物对大脑影响的另一些变量产生的。相同的DNA数

据还可能发现影响个体对一定的药物副作用的易感性的基因变量。所有这些遗传信息将指导我们针对不同患者选择不同的治疗方法。

DNA数据还能帮助我们重新确定不同精神疾病的界线，它们通常是相互重叠的。当然，它也能帮助我们确定所谓正常和病态的行为模式的界线。把基因变量与脑功能研究、正常心理测试和个人的生活经历结合起来，就有可能以每个患者的个性特征来取代粗略的诊断标准。

50年后，我们会同今天一样需要心理咨询。有些患者会残废，也许因为觉得自己无用，或者幻想自己无所不能；也许因为莫名的恐慌；也许因为大脑里回响着恐怖的声音。另一些患者可能觉得没有快乐，没有生气，悲观失望，无限焦虑。还有些患者可能只想清点自己的生活。

不过在50年后，去看心理医生的人都将带着一个新的信息源——一个密码，它能使我们进入国家卫生部门计算机的个人DNA文件。在那个文件里存着每个人的基因序列，并且特别提醒需要注意的个人基因变量和组合，它们可能影响一系列不同疾病的易感性，或者影响药物的功效。

第一次咨询可能需要一个小时。三分之一的时间用于常规问答，如个人情况、家庭背景、特殊症状。其余的时间用来进行信息交流。最后，心理医生会拿出一个意见，建议做一些诊断测试，还要求打开患者的DNA文件。

这样的要求在那时不会有什么麻烦。依法建立起来的DNA文件的国家数据库，确保了个人隐私，同时也为专业人员用它来造福社会提供了条件。许多访问心理医生的人都是愿意答应的。当患者来自有某种精神疾病的家庭时，这将有特别的意义：他想知道，自己患病的机会有多大，有没有什么预防的办法。寻医找药的人，会在他们特殊的基因变量组合的引导下决定他的去向。

这样的引导有着特别重要的意义，因为我们那时可能需要在几百种治疗方法中做出抉择。有些方法我们现在已经有了，但将来会更好，对神经传递作用起着更具针对性的作用。另一些将在我们今天关于大脑功能的新认识的基础上发展起来。还有一些则需要等着我们把增大精神病风险的基因变量找出来。

精神疾病的遗传信息不但能改变精神病医生的诊断和治疗，精神病学也将因此而更多影响我们对自身的思考。在20世纪的前50年里，精神病学曾帮助我们发现，我们都受着固有情绪的强烈影响，并从对它们的认识中获得好处。在后50年里，精神病学给我们带来了抑制疯狂行为的药物，说明了我们是如何依赖于像5-羟色胺和多巴胺那样简单的大脑化学物质。影响不同行为模式的基因变量的发现，为每个人的独特的生物学填充了一些重要的事实。在这些基因变量中，尽管可能有很多难以解释的，但一定有一些会成为有用的工具，我们可以拿它们来思索和构造我们自己的生命故事。

巴龙德斯（**Samuel Barondes**）

巴龙德斯（Samuel Barondes）医学博士，旧金山加州大学神经生物学和精神病学中心主任，Jeanne和Sanford Robertson讲座教授。他还是国家精神健康研究所科学咨询委员会主席。他的作品有《分子与精神疾病》（Molecules and Mental Illness）和《情绪基因：追寻癫狂抑郁之源》（Mood Genes: Hunting for Origins of Mania and Depression），现在正写一本关于精神病药物的书。

心理疗法和穿在身上的电脑

N.艾特柯芙

Nancy Etcoff

哲学家布伯（Martin Buber）在自传《我的哈西德之路》（*My Way toHasidism*）中[118]，讲了他1910年在布科维纳遇到的一件事情。那时他刚做完演讲，正在一家咖啡屋休息，一个中年男子（他只知道是M先生）向他走过来：

> “博士，”他说，“我有一个女儿。”停了一会儿，他接着说，“还有，还有一个小伙子。”他又停了一会儿，“他是学法律的学生。他以优异的成绩通过了考试。”又停了一会儿，时间更

长……“博士，”他问，“那人可靠吗？”我很惊讶，但觉得不好拒绝他。“哦，M先生，”我向他解释，“从你的话，当然可以认为他是一个刻苦能干的人。”他又进一步问，“可是博士，”他说，“他也有好的头脑吗？”

“那就更难说了，”我回答说，“但不管怎么说，仅凭刻苦是不会成功的，他还需要一定的头脑。”M先生又停了一会儿，然后问，显然是最后一个问题了：“博士，他现在该做法官还是律师呢？”

“这件事情我没有什么可说的，”我回答说，“我不认识那个年轻人，实际上，我从来没有听说过他，在这个问题上我提不出什么建议来。”M先生用几乎悲哀失望的眼光看着我，一半抱怨，一半理解。接着，他用难以言状的语气，无奈而谦卑地说，“博士，您是不想说。”

每个星期有两个下午，当我离开有序的世界，走进疯狂隐秘的心理治疗的世界时，常常会想起布伯的故事。我跟病人的对话充满了戏剧性，幽默而痛苦。他们有问不完的问题：

“我是不是应该离开妻子？”

“我是不是性迷恋者？”

“我为什么老觉得疲惫？”

“我是不是该停止跟我兄弟说话？他总是冒犯我。”

“我怎么知道自己会不会因为触摸音像店里的CD而被污染呢？”

“我怎么才能确信不会伤害我的小宝宝？”

因为这些问题，我想回到实验室，做一些计算，描绘某些更基本的大脑功能区域。它们也令我感到无能，想知道我们的谈话怎么才能有所帮助。

21世纪到来时，我们比以往任何时候都更多地认识了大脑的网络、

思想的计算和隐藏在人类基因组里的密码。尽管把这些研究转化为有效的治疗方法大有好处，但大多数的进步还有待于未来。神经科学得意忘形的时候，是精神病学垂头丧气的时候，也是临床心理学谨慎乐观的时候。如果目前的趋势发展下去，50年后就几乎不会有精神病医生了。医学院里选择精神病学专业的学生比1929年以来的任何时候都少了。调查结果说明了原因：他们觉得精神病学对病人没有多少帮助，在智力上也没有多大的挑战性，声望不高，而且报酬比其他专业的少。

不过，潜在的病人最好还是清醒一些。在群体调查中确诊有精神问题（根据《精神障碍诊断与统计手册》第四版确定的诊断标准）的大多数人并不接受治疗。有些人没有把他们的症状与精神障碍联系起来。许多人说，他们能靠自己解决问题——或者靠家人和朋友的帮助，或者靠祈祷、休息、锻炼，靠维生素、镇痛药、烈性酒。有的人没有医疗保险，有的人感到羞怯或者害怕担上精神病的名声。不过有一项调查提出了不同的问题：大约一半受调查者报告说，他们不太相信标准的精神病治疗方法，如药物和心理治疗。精神治疗医师并不缺少病人，但是在向他们求助的人当中，许多人只是“疑病”[119]——他们不能被诊断为任何精神病。

这个问题还将变得更加尖锐。德斯加莱斯（Robert Desjarlais）和哈佛大学社会医学系的其他人员在他们的报告《世界精神健康：低收入国家的问题和任务》（1995）中预言，患精神障碍的人数将在全球迅速增长，部分原因就是人的寿命更长了，活到一定的年龄时，患某些疾病的风险也大了。到2020年，抑郁症将成为仅次于缺血性心脏病的第二大致残因素。对日益增多的抑郁，人们把原因归结为我们周围发生的所有事情，例如与社会的隔绝和社会角色的失败，全球食物结构的改变（如低Ω-3脂肪酸），诊断标准和评价方法的改变以及许多虚报的抑郁患者人数（因为像氟西汀那样的药很难卖给精神病医生和病人）。

我不是未来学家，不过我还是想为21世纪的心理疗法做些预言。当前的不满是一个很好的信号，它预示着全方位的变化。我从我认为不可避免的理论重心的转移说起，最后就未来可能的精神疗法谈几点预言。当然，我还要说，即使未来的氟西汀很好用了，人们还是愿意谈话的。

知识的融合

有一点很容易预言：把先天遗传或后天环境作为绝对唯一的影响因素的思想，跟其他无用的老观念一样，将被扔进垃圾堆。精神病学的问题不可能是单个基因或单个神经传递介质（如5-羟色胺、多巴胺等）引发的。看见“原初场景”或者发现女孩没有阴茎，都可能引发精神问题。[120] 大多数障碍的根源都在于基因与“环境”的复杂的相互作用——这里的环境覆盖了所有的非基因因子，也包括随机作用。作为概率风险因子发生作用的多个基因很可能会影响大多数的精神障碍。

还有一点进步，虽不那么显而易见却是不可阻挡的：精神病学家们如果还认为大脑与他们的作为无关，就要惹麻烦了。50年后，思维与大脑的研究不会像今天这样，属于分离的研究部门或专业。19世纪精神病学与神经学之间发生了恶意的领地争端，结果，大脑及其“器官”和“神经”的障碍划给了神经学，思维及其“功能”和“精神”的障碍划给了精神病学。不过，所有的精神过程当然都发源于大脑的计算，于是思维和大脑的研究都是一个连绵的知识整体的组成部分。

对于那些不能想象与神经科学亲密共处的心理分析学家和人道主义临床医学家，我请他们想想人的大脑——从普通的眼光看，它是公认的缺乏美感的东西；然而走近来考察，它却美妙无比。这个1千克多重的器官——包裹着数十亿神经元，同银河系里的星星一样多；20万个突触联络把那些神经元彼此联系在一起——是宇宙间最复杂的结构。有的科学家用脑图像仪器来观察记忆、想象和欲望时刻的大脑，在他们看来，大脑是令人敬畏的。但是依然存在一个大问题：血液的涨落和复杂的联系网络如何成为我们感觉的经验和我们思想的内容？这正是未来50年将要困扰我们的问题。正如遗传学家雅各布写的， [121] “即将结束的世纪已经被核酸和蛋白质抢占了。下一个世纪关心的是记忆和欲望。它能回答它们提出的问题吗？”

脑科学与心理治疗的实践有什么关系呢？有人说过（最强有力的是神经生物学家坎德尔[122]），心理治疗不仅改变我们的思想，还改变我们的大脑——真的。有效治疗发生作用的方式和机制，跟其他任何形式的强化学习是一样的：它改变基因的表达，从而改变突触联络的强度并产生改变大脑神经细胞联络模式的结构变化。可能有人觉得这跟专业音乐家的训练有些相似。神经学家帕斯库尔-利昂（Alvaro Pascual—Leone）指出，专业音乐家的大脑在训练时会产生功能和结构的改变，那些改变可以通过神经成像技术来证实。他进一步提出，激烈的智力演

练也都可能产生这样的改变。

越来越多的研究通过观察治疗前后的脑图像来对比精神治疗与药物（如氟西汀）治疗的效果。这些研究已经针对OCD（强迫症）和严重的抑郁症做过了。研究发现，当两种治疗方法都有效时，它们会产生相同的大脑变化。研究似乎说明复杂的心理学变化最终有着相同的路径。将来，只要扫描病人的大脑，我们就可能解决表面看来不可能解决的难题：如何确定治疗是否有效？该在什么时候终止治疗？

弗洛伊德走了，达尔文来了

在弗洛伊德理论的影响下，从20世纪40年代至70年代（有的甚至直到今天），心理治疗专家都想当然地认为，精神疾病源于很早的儿童时代，因此在治疗中必须重构遥远的过去——可是有限的记忆担不起那么沉重的任务。阿伦（Woody Allen）在影片《安妮霍尔》里通过辛格（Alvie Singer）这个角色，嘲笑了他15年的分析。[123] “我想再做一年，我也该到洛德斯（Lourdes）去了。[124] ”精神分析学家们借了一句话（“你不可能用勺子挖通苏伊士运河”）来说明那任务确实太艰巨，需要很长的时间。但研究并没有证实儿童时代的创伤是任何精神障碍的根源；即使创伤后压力障碍（PTSD）也可能源于长大以后的经历。虽然微观分析病人的早年经历也许能揭示某些东西，但今天的认知科学却说明记忆是可以塑造的。大脑就是通过遗忘、阻滞和偶然的错误记忆在运行的，我们在一般情况下服务良好的记忆系统，都要付出这样的代价。

神经病学家斯通（Alan Stone）在他1997年的一篇文章《哪里还将留下精神分析？》中得出这样的结论：“不论作为理论还是实践，精神分析都是属于人文而不属于科学的艺术形式。它距离文学比科学更近。”如果说弗洛伊德让栖居变成人文和艺术，那么达尔文则走进了行为科学和医学。50年后，达尔文的医学将为这个领域搭起一个框架。大脑和其他人体器官一样，在自然选择中形成，然后演化出精神的模块，增强繁衍的适应力，从而保证生存和延续。心理治疗实践的中心将从疾病转向疾病的易感性，从症兆转向适应性反应，从单纯的个体的病史转向人类共同的经验。

某些精神疾病的流行可以追溯到诱病基因所带来的健康好处。躁狂抑郁症就是一个例子：与轻度躁狂相关的活力、创造力和魅力，能为某

些得病的人，或者虽然没有得病但基因仍然起着有益作用的人，带来健康优势。其他疾病可以理解为大脑的某些模块出现了问题——模块也就是系统，例如（在精神分裂中）通常区别我们自己与别人的行为的系统，或者（在孤独症）使我们理解别人的倾向和感觉的系统。某些症兆可能是为了平衡现代环境与祖先环境之间的不对称，也可能只是过分的正常反应。例如，我们知道的一般性焦虑可能演化为对不确定危险的反应，而恐惧症则可能演化为对特定危险（如流血、登高和毒蛇）的反应。PTSD患者常常回忆起那可怕的一幕，这也许是痛苦的，但它能唤起人们警惕威胁生命的危险，从而在未来避开那些危险。轻度抑郁可能起着适应的作用，在困苦的时候积蓄力量，在需要帮助的时候向别人发出信号，在安静的时候能重新评价自己的目标。当个人不能或不愿与他周围的人对立时，轻度抑郁也可能是屈服的信号。

从进化的观点看，悲伤、恐惧、愤怒、厌恶、羞怯和愧疚都是为了适应和防御；就像咳嗽、老茧和疼痛，都起着良好的作用。它们也可能是非常有害的，因为它们的作用服从密歇根大学精神病学家内斯（Randolph Nesse）所谓的“烟火探测器原理”：错误警报总比没有警报好。我们的环境比祖先的环境安全多了——我们远离了传染病、营养不良、吃人的野兽和自然的灾难。也许，有讽刺意味的是，今天处在焦虑重压下的人们更能适应遥远过去的环境。

进化论的解释提出了有关治疗的重要问题。假如某些症兆（如恐惧）起源于生物学，是不是意味着不可治疗呢？不是那样的。对血、蛇或其他动物的恐惧，经过几个小时的暴露治疗[125]都能消除。不过，虽然强迫性障碍、抑郁、惊慌和恐惧有着明显的危害，能通过治疗来缓解，但轻度的焦虑和抑郁，却可能最终是有用的。这些痛苦的状态能帮助我们改变生命的过程，怀疑自己或他人的决定，与朋友和家庭和睦相处，避免危险。正如内斯说的，通过药物来增强社会的免疫力，使它没有悲伤和失落的感觉，也许不是什么好事。

在另一个范式转移中，健康的研究将变得与疾病的研究一样重要。我们将看到，在从实证心理学到分子遗传学的诸多领域里，科学家们研究什么在危难的时刻保护我们，什么在为我们抵御紧张，什么基因、环境或性情可能影响健康。“精神健康”不再是错误的名词，因为它的领域将不仅仅是研究疾病。

从心理治疗到可穿戴的电脑

传统的精神动力学治疗的核心是患者与医生之间的关系。弗洛伊德唤起我们文学的想象：精神分析的对话贯穿了人的一生，开创了令人满意的意味深远的人物传记。正如斯特拉奇（Lytton Strachey）在《维多利亚名人传》（*Eminent Victorians*）的前言中写的，传记是“一切文艺分支里最能深入人性的”。[126] 但是，即使没有保健组织（HMO）的帮助，也一定有别的手段来取代悠闲的精神分析法。人们批评精神分析是“用于不确定问题的没有确定结果的不确定技术”，对大多数在精神重压下的病人来说，它从来就没有意义。

谈话治疗的未来如何？对话的内容是什么？多数时候，心理治疗都以问题为中心，在手册的指导下进行。它很简短，立足当前的状况，依赖于过去证明对某些问题有效的技术。做这种治疗的，可能是心理学家（他们将来会有处方权）或社会学工作者，或者偶尔还有精神病学家。医生的热情和关心很重要，不过治疗不在于关系，而更多地在于信息的交流。治疗的场所可能随时变化，不一定总是面对面的。越来越多的治疗可能会远距离实现，例如通过教育、自我诊断和治疗监测的互联网；通过掌上向导（Palm Pilot），它能告诉我们在恐怖袭击的时候应该做什么；还可以通过戴在身上的器件。心理治疗将不再只有谈话而没有行动。在长期太空旅行的计划中，美国国家航空航天局（NASA）正投资制造能发出抑郁、焦虑或疲惫信号的服装。正如宇航员里莱乌闵（Valery Ryumin）说的，“把两个男人关在一个仓里，让他们在一起生活两个月，就可能出人命。”[127] 预警响应工具能预报宇航员应该什么时候停止工作，什么时候接受不同的治疗，告诉他“听你的认知行为治疗磁带”“服抗抑郁药”，或者“休息一会儿，三个小时以后再来测试”。地面的患者想退出治疗，也可以通过这种工具的信号来监测自己是不是旧病复发了。

未来的计算机也许能识别我们的情绪。固定在我们的衣服、首饰或者眼镜上的那些小机器，可以通过我们想象不到的参数来监测我们（例如，拿眨眼的次数或额头上的一道皱纹来跟我们正常的基本数据进行对比），而植入体内的器件可以监测我们内部的活动。麻省理工学院（MIT）媒体实验室的皮卡德（Roz Picard）1998年在《大西洋月刊》的一次访谈中提出，情绪传感的可穿戴电脑将“提取你表面的气味和隐藏的气息……它们将跟内衣一样，不可能与人共享，从而成为真正的个

人计算机。”这样的话，它们也许更像戴在身上的临床医生，而不像女人的内衣。[128]

将来我们可以喝药物的鸡尾酒，为什么还为心理治疗烦恼呢？假如药物能减轻令人残疾的顽症，多数人当然愿意选择它们。但是多数研究表明，药物对一些人有效，心理治疗对另一些人有效，而两者的结合才是最好的。药物控制征兆，而心理治疗帮助人们解决问题，学会解决方法——更不用说人们更喜欢在心理治疗的同时继续服用药物。对许多人来说，单纯的心理治疗还是最好的选择，它跟药物治疗一样对大脑产生作用，而且几乎一样快，却不需要花多少钱，也没有副作用或潜在的危险。另外，药物似乎只有在服药的时候起作用，而谈话治疗能通过学习保持长久的影响。

有时候，未来的治疗也可能很像传统的治疗。缺乏面对面交流的简单治疗可能对某些人起不了什么作用——也许正是那些在今天没能从心理治疗得到好处而需要长期甚至永久治疗的人。对他们来说，真正有效的治疗应该是联合一个临床医生，一个能帮助他们过安定生活的医生——关心他们，认真听他们的话，用心与他们交谈，为他们提供一个安全的港湾，让他们走出孤独的生活，使他们感觉到一个共同的目标。在未来的世界里，精神恍惚的人们游荡着、寻求短暂的归宿；到处可以看到大脑的扫描和大脑的激发，然而改变我们认识世界的方式的对话，还将继续进行下去。

艾特柯芙（**Nancy Etcoff**）

艾特柯芙（Nancy Etcoff）是哈佛大学医学院、麻萨诸塞州普通医院精神病学科和哈佛“思维-大脑-行为行动”成员。艾特柯芙博士对美、情绪和人类面部的研究曾发表在《自然》《认知》《神经元》等科学杂志，被普及读物广泛引用，赢得过许多奖励。她也是《漂亮者生存》（Survival of Prettiest: The Science of Beauty）[129] 的作者。

征服疾病

P. W.埃沃德

Paul W. Ewald

疾病是什么引起的？这个问题太基本了，现代技术那么复杂，专家的承诺又那么自信而超前，外行的人一定以为疾病的起因已经得到了很好的认识。不是的。为了把握许多吞噬生命的慢性疾病的起因，如心脏病、卒中、阿尔茨海默症、精神分裂、癌症和糖尿病，今天的医学还在努力奋斗着。在未来50年，我们的生活质量将依赖于我们能在多大程度上控制这些慢性疾病。

跟所有生命现象的情形一样，我们可以在机械论的意义上（“产生疾病的病原体是什么？”），也可以在进化论的意义上（“导致疾病的选择性压力是什么？”），考虑起因的问题。在机械论的意义，教科书上罗列的那些疾病，大约只有一半的起因形成了共识。机械论的起因可以分成三种：遗传的影响、寄生（包括感染）的影响和非寄生的环境影响（如辐射、过多或过少的特殊的化学物质）。

健康科学的多数专家都提倡以一种积木式 [130] 的方法来解决起因问题。他们试图通过在细胞和生物化学的水平上认识疾病的行为，从而获得最后的解决。这是很诱人的想法，但还没有产生任何重大的医学进步——重大的进步指的是决定性的解决，而不是零碎的修补。托马斯（Lewis Thomas）30年前在“医疗技术”一文里讲的话，今天仍然是正确的：医学实践的绝大多数，从器官移植和搭桥手术到多数癌症的治疗，都是些修补的解决办法，或者只是意义不大的支持性治疗。 [131] 专家们认为，今天的病不像过去那样好治，我们所能希望的最好结果大概也只有修补了。不过事实似乎不是那样的。对一些最近认为不可救药的疾病，我们还会继续发现一些决定性的解决办法。这些方法一般依赖于我们对感染原因的认识。例如，在过去20年，通过检测输血者的乙型和丙型肝炎病毒，通过使用乙肝疫苗，成千上万的人躲过了肝癌和肝炎的危害；胃溃疡和胃癌现在也能用抗生素来治疗。

不过，即使在感染性疾病中， [132] 基础性的成就也更多地来自演绎推理，而不是积木式的归纳法。詹纳（Edward Jenner）在200多年前建立现代接种疫苗时，还没有认识病毒； [133] 半个世纪以后，斯诺（John Snow）和西莫威斯（Ignaz Semmelweis）建立现代流行病学时，从没见

过细菌，他们说明的是感染性疾病如何传播，如何通过阻断传播来预防。甚至李斯特（Joseph Lister）在证明手术器械杀菌作用的时候，也没有见过细菌。埃尔利希（Paul Ehrlich）和弗莱明（Alexander Fleming）在建立抗生疗法的概念、实践抗生疗法的时候，根本不知道化学物质是通过什么机制来抑制细菌生长的。[134] 这段历史说明，为了解决可能产生系列决定性突破的概念问题，我们需要超越指导现代医学的积木方法。途径之一，就是以进化的观点来认识疾病的因果关系。

进化论的解释最终建立在遗传因果的基础上。于是，我们可以以为，进化论医学主要是在人类遗传的背景下解释人类疾病。进化论医学确实在做这样的解释。例如，慢性病就是用衰老理论来解释的：自然选择倾向那样的特征，年轻时它们有利于个体成长，年老时它们成为负担，自然选择的作用也弱了，因而身体也崩溃了。另外，慢性病也可能源于环境的不对称：现代的环境跟过去决定人类基因形成的环境是大不一样的。第三个主要的原因是感染：慢性病可能是因为某些感染性病原体隐蔽地破坏了组织，最后通过一系列严重的疾病表现出来，如心脏病、癌症或阿尔茨海默症。这种可能性不会否定最终的基因背景下的进化论解释的意义，不过它提醒我们不能把慢性病的因果假定完全限定在人类的基因，还必须考虑寄生者的基因。

我的同事科克兰（Gregory Cochran）和我曾经强调，对具有普遍破坏性的慢性病来说，在进化论观点下考虑的证据中隐含着感染的作用。假如这些病都是有缺陷的等位基因引发的，那么基因的突变率一般应该很低，这样疾病发生的频率才可能总是我们看到的那样。人类生存和疾病发生的模式并不满足衰老的“取舍（trade—off）”模型。尽管在原则上新的环境因素也是可能的，而且在某些情形还很重要（例如吸烟导致的肺癌），但是，非感染性的环境作用一般不能充分说明疾病发生的模式。把感染的作用考虑进来，我们就能很好解释大量记录在案的疾病流行模式，这样做也符合进化论的原理。因为，宿主和寄生的病原体之间基因的利害冲突，可以将疾病状态无限期地保持下去。在短时间里，对某种特殊病原体的遗传易感性会像坏基因一样被自然选择淘汰。不过，跟坏基因不同的是，在长时间里，对感染性疾病的遗传易感性在协同进化的“博弈游戏”中不断变化着。[135] 自然选择淘汰了对病原体的遗传易感性，留下的是有办法抵御那个病原体的宿主群。这些办法反过来又促进了病原体群的对抗方式的演化，游戏就这样无限进行下去。

进化生物学的许多问题解决起来都很慢，因为问题太复杂，试验很困难，还缺乏经费。相反，慢性病的原因也许能较快得到解决——不是因为健康科学的传统学科的研究者会急着根据进化论来考虑他们的问题，而是因为他们眼下正在争论，在热情地研究慢性病的感染原因。有了进化论的背景，有了医学研究的进步和积累的证据，我想，在未来50年，大家会相信那些常见的高危慢性疾病——精神分裂症、糖尿病、阿尔茨海默症、大多数癌症和大多数生殖问题——都是感染引发的。

可是这个预言是相当软弱的。我大概等不到能评判它的时候让别人来批评我错了。如果进化论的方法对疾病的起因真有意义，我们应该能够担起挑战的风险，去确定哪些疾病在未来几十年可能被认为是感染的结果。下文中的表列举了最重要的一些慢性病，还注明了年代，我猜想到那个年代我们可以接受它们是感染引发的。为了能检验我的预言，我提出一个“接受”准则：那个年代到来前5年出版的医学教科书，至少有四分之三会把它们作为感染引发的疾病。

如果假定人们主要是凭证据的质量来接受哪些疾病是感染引发的，那就等于过分加强客观评价的分量。只要我们在恰当的范围内检验假说，证据越多当然越有帮助。但我们也许永远不会得到顽固的反对者们要求的那种感染证据，因为对残存的大多数感染引发的慢性病来说，不可能获得感染原因的绝对证明。例如，在过去的四分之一世纪里，医学界已经公认疱疹病毒8引发了卡波西（Kaposi）肉瘤[136]，人类T细胞白血病病毒（HTLV—1）引发了成人T细胞白血病。但同样质量的相关证据对动脉硬化症来说却是不能接受的，主要是因为多年的研究和专家对那种病的意见形成了不同的既得利益，阻挠了接受的步伐。摆脱这些阻碍还需要时间。这应了达尔文、普朗克（Max Planck）和库恩（Thomas Kuhn）的论断：相当比例的老卫道士们将不得不退却或离开，相当数量的没有任何既得利益的年轻人正走上舞台，他们一定会成熟起来，产生影响，打破专家们的意识的平衡。我想，到2015年时，在动脉硬化和阿尔茨海默症问题上可能会这样。这里还有一点黑色幽默的东西：病原体在这个过程的作用是，谁驱除它们，它们就驱除谁。

表　不同慢性病的感染因果认定时间预测

疾病名称	可能的病原体	接受时间
多发性硬化	肺炎衣原体，人疱疹6	2010
2型糖尿病	丙型肝炎病毒（次要原因）	2010
2型糖尿病	未知（主要原因）	2025
头颈部肿瘤	人乳头瘤病毒	2010
鼻咽癌	EB（Epstein Barr）病毒	2010
小儿白血病	未知	2015
乳腺癌	小鼠乳腺病毒，EB病毒	2015
动脉硬化症	肺炎衣原体，牙龈卟啉单胞菌，巨细胞病毒，放线共生放线杆菌	2015
阿尔茨海默症	单纯疱疹1，肺炎衣原体	2015
精神分裂症	弓形虫，单纯疱疹2，内生反转录病毒，波纳病病毒（BDV）	2025
双相抑郁症	BDV	2025
前列腺癌	未知的反转录病毒	2025

表中的预言以过去的转变为指南，但是没有简单地认为从第一个令人信服的证据出现到人们接受它，会经历相同的时间。另一个考虑是我们从现在接受的感染性疾病出发会跨出多大的一步。我想，某些头颈部肿瘤的感染原因可能会比动脉硬化症的感染原因更快地为人们所接受，尽管现在关于前者的证据更不完整。我们优先考虑的是主要原因。人们已经相信，头颈部肿瘤的可能病原体（人乳头瘤病毒）是宫颈肿瘤的原因；而且，关于那些癌症（如成人T细胞白血病和卡波西肉瘤）的均衡考虑，已经没有强大的既得利益的障碍了。同样，假如将来认为肺炎衣原体是隐藏它的某种主要慢性病的原因——动脉硬化症、中风、阿尔茨海默症和多发性硬化——那么，反对它作为其他疾病的原因的障碍，就会像多米诺骨牌一样坍塌下去。不过，在这种情形，共性不在于组织型，而在于一种共同的诱病基因——ε-4等位基因，它是跟增大的肺炎衣原体的易感性联系在一起的。

人们似乎普遍地特别顽固地反对癌症的感染原因。从实际的证据看，这些阻力令人惊奇，因为医学研究者们已经承认，大约15%到20%的人类癌症（25年前还不到1%）是感染的结果，而没有几个癌症病例（不到5%）能排除感染的因素。1910年，劳斯（Peyton Rous）证明病毒可以在小鸡身上引发癌症[137]，大约那时以来，人们就在争论人类癌症的感染原因。20世纪70年代末，感染原因的最坚决反对者们宣扬他们胜利了，那时确认了致癌基因和非感染性致癌物质的作用，从而获得了一个可能的机制。他们还在继续宣扬胜利，不过人数一年比一年少了。他们犯了一个逻辑错误，错把支持一个机制的证据当成反对另一个（与它相容的）机制的武器。疾病的这三种原因经常是共同发生作用的。

跟动脉硬化和癌症的情形一样，严重精神病（如精神分裂和双相情感障碍）的感染原因，由于专家们的利益争夺，肯定也要拖到很久以后才能被接受。除了这一点，精神疾病是人类特有的，很难为感染原因做出实验的证明。我们怎么知道小老鼠什么时候在幻想还是患了偏执狂、抑郁症或躁狂症？为了反驳这种观点，可能有人会提出我们以前关于梅毒患者的精神错乱的认识。不过，它的感染原因之所以容易被人接受，是因为研究者们早就把那种精神错乱跟梅毒联系起来了。一旦承认了梅毒是感染引发的，专家们自然会承认患者的精神错乱也是感染引发的。于是，梅毒患者的精神错乱被当然地隔绝在精神病范围内一个孤立的感染因果的圈子里。今天那些原因不明的精神病还没有与能清楚认定的急性病联系起来。在这个意义上，它们更像成人T细胞白血病和卡波西肉瘤，而不大像梅毒患者的精神错乱。既然这样，我为什么还那么乐观地认为精神分裂的感染原因会很快得到承认呢？流行病模式明显牵扯着感染的原因。例如，精神分裂病人似乎多出生在暮冬和早春，这意味着精神分裂症可能是在出生前或刚出生的时候，由暮冬和早春盛行的某些感染性疾病引起的。我们还发现它们与某些可能的病原体有着重要的关系，这也牵扯到感染的原因。例如，最近的研究表明，在精神分裂症患者中，42%的人对寄生在大脑的刚地弓形虫呈阳性反应，而在对照实验的健康人群中，只有大约11%。10年前不承认精神分裂症的感染原因的专家，现在也认真看待这种观点了，有几个实验室正在积极研究这种可能性。在我看来，所有这一切都意味着，在未来10到50年里，旧势力的退却和离开，一定能打破过去的认识平衡，在很大程度上使我们接受精神分裂的感染原因。

在未来50年，紧跟关于病原体的发现，我们将针对感染引发的慢性

病开发疫苗。假如一个病原体不是特别易变的，疫苗就会特别有效。于是我们可以预料，DNA病毒（如人乳头瘤病毒）会比RNA病毒（如艾滋病病毒HIV）更容易得到控制。将来像这样去把握疾病，不需要开发什么根本性的新技术。我们知道，恰当的疫苗能预防疾病，正确的抗感染作用能治疗疾病，防止感染性疾病的流行和危害。我们已经知道，中断传播链可以防止个人染病，有时还能防止疾病在整个人群中传播。为了以这样的方法解决新发现的感染因子，药物跟踪记录是行之有效的。

对感染性疾病来说，进化的原因主要关系着致病力的演变：寄生的病原体因为什么而演变为良性的、恶性的或居于两者之间的？问题的答案有望为我们带来第三种控制感染性疾病的方法。那就是，除了治疗感染、抑制感染的扩散，我们还应该能控制病原体的演化，从而把威胁生命的恶魔转化为友好相处的伴侣。理论和现有的证据说明，对每一类主要的疾病，至少有一种干预的办法能做到这一点。我想，在接下来的四分之一世纪里，我们能看到第一个成功演化的例子——例如，也许是通过疫苗的巧妙运用，也许是在贫困国家阻止腹泻病原体的水路传播。还有一个悬而未决的问题是对抗生素抗性的控制：越是良性的病原体，抗生素用得越少，抗生素抗性的演化也越弱。如果我们能控制致病力的演化，那么我们也应该能控制抗生素抗性的演化。

如果说医学的历史是一个指南，那么，第一次成功将激励我们把同样的方法用于其他的疾病，从而也取得像疫苗、抗生素和卫生保健那样的进步。然而，因为很难预测第一次成功什么时候出现，所以后来进程的时间也很难预测。到世纪中叶的时候，我们也许能使几种疾病，例如某些破坏性的腹泻，从恶性的转化为良性的。但对其余大多数疾病来说，我们仍将处在试验的迷雾里。

哪些事情不会在未来50年发生呢？我们朝着决定性控制的努力肯定会遭遇一些阻力。例如，除非出现什么根本的新方法，HIV仍然是一个困难重重的问题；修补式的方法能抑制它却不能解决它——就是说，抗病毒药物和疫苗只能抑制它，却不能决定性地控制它。这种悲凉的前景，源自HIV的遗传适应性：它很容易产生对抗病毒药物的抗性，而且很可能逃脱疫苗的打击。像过去的几年一样，巧妙结合各种抗病毒作用可以提高疗效，但不能根治疾病。我们所能希望的只能是争取时间，延续生命——从抗病毒药物那里争取10年左右没有艾滋病的生活；在感染后立即使用治疗疫苗，大概也能争取10年；通过控制致病力的演化，大

概还能争取10年。这几个十年虽然都有意义，但还是不能像我们用药物控制白喉、天花、百日咳、麻疹和脑灰质炎那样，决定性地控制艾滋病——特别是，抗病毒化合物的长久使用，会产生许多可怕的副作用。

尽管我们现在担心会突然出现什么新的可怕的席卷全球的疾病，我们还是有理由相信，在未来的半个世纪，像艾滋病那样的毁灭性疾病的恐慌，也许没有，也许有一个——很可能没有。一个多世纪以来，地球上相隔遥远的人，比以往任何时候都更快地融合在一起了。人的融合必然使我们在今天面临着几乎来自世界各个角落的病原体。不过，在这样的大融合中，我们现在知道只有一种新的病原体，原来局限在一个偏远的地方，然后扩散开来，导致了可怕的全球流行病：那就是带来艾滋病的人体免疫缺陷病毒（HIV—1）。通过艾滋病流行的例子，科学家和非科学家都在警告我们，其他病原体也许会产生同样的威胁。在那些可能的病原体中，我们熟悉的有奇异出血性病毒（如埃波拉（Ebola）病毒和马堡（Marburg）病毒）；蚊子携带的西尼罗河病毒；汉塔病毒（最近，它在美国西南部引发了“四角”病）；引发疯牛病和相关人类疾病的朊病毒（prion），新发现的变异克雅氏（Creutzfeld—Jakob）病毒。[138] 不过这些忧虑都是多余的。这些病原体虽然对病人来说很可怕，但不会产生全球性的危害。它们完全不具备那样的特性。西尼罗河病毒不可能从人向蚊子传播。埃波拉病毒在外部环境下不能长期生存，不可能在人群中形成致命的感染链。朊病毒只有通过非常的而且可以控制的行为（如食人肉和角膜移植）才能在人群中传播。

真正的威胁来自那些已经在我们身边的病原体，有的在吞噬我们的生命，有的可能还会增强致命的能力。假如现在牵涉疾病感染原因的证据是可靠的，那么我们正在因为可怕的全球流行的疾病而失去自己的生命，如心脏病、中风、阿尔茨海默症和癌症——它们都源于要命的然而却被我们忽略了的感染。我们一直在为几只迷路的小猫担心，却不知道自己的身后跟着美洲豹。在未来的半个世纪，我们将最终发现那些美洲豹，而且至少有办法阻击其中的几只。

埃沃德（**Paul W.Ewald**）

埃沃德（Paul W.Ewald）是Amherst学院生物学教授，他帮助创立了进化论医学，为它在世界许多大学做过专题演讲、讨论班和学术会议。他是《感染性疾病的进化》（Evolution of

Infectious Disease ）的作者（被誉为本学科出现的分水岭），还写过一本《瘟疫时代：隐藏的感染如何引发癌症、心脏病和其他致命疾病》（Plague Time: How Stealth Infections Are Causing Cancers, Heart Disease, and Other Deadly Ailments ）。

译后记

译者

2003年8月25日

海南岛上，科罗旺台风下

20世纪的教育似乎自然而然地分化为人文学科与自然科学，20世纪的文化也就是C.P.Snow 1959年在同名著作里概括的“两种文化”：文艺家和科学家。他惊奇地发现，20世纪30年代的人文学者们喜欢说自己是“知识分子”，仿佛没有别的人了，连冯·诺伊曼、爱丁顿、爱因斯坦、玻尔等也都不算。原因是科学家没有很好地让人们了解他们的工作，没有成为报刊杂志的主角。1963年，在《两种文化》的第二版里，Snow猜想会出现“第三文化”，两个文化间的距离会拉近，文艺家能与科学家对话。近年来，许多科学事情确乎发生在人文学者的头上，所以，John　Brockman认为真的出现了第三种文化，具体表现就是科学家为大众写他们的发现，并且跟我们的现实世界联系起来。过去属于哲学家和诗人的领地，如今被科学家抢占去了。科学家眼光向前，改变世界，而哲学家和诗人还跟在他们前辈的身后。1991年，Brockman借Snow的“第三文化”的名词，表达了不同的意思。在他看来，“文艺家没有跟科学家交流，而是科学家直接与大众交流。传统的知识媒体做的是流水作业：知识分子写（write down），记者编（write up）；第三文化的思想者们不充当中间角色，而是以能为广大读者所接受的方式表达他们深刻的思想。”在引言里，Brockman重复了自己的定义，第三文化“由那些科学家和经验世界的其他思想家组成，他们通过自己的工作和通俗的作品，正在取代传统知识分子的地位，让我们更真切地感觉生活的意义，重新认识我们是谁、是什么。”他主持的“边缘网”（www.edge.org）就是第三文化的一个舞台（有关第三文化的材料和那些科学家的讨论，都可以在那个网站看到），读者面前的这本书，就是25个第三文化的精英们联合演出的一出多彩的大戏。

这是一出传统剧目，我们已经看过了许多。在新世纪的起点，“我们谁不想揭开未来世界的帷幕，窥探下个世纪科学发展的奥秘？”1900年，希尔伯特向未来数学提出23个问题的时候，开口说的就是这句话；百年以后，又一个大数学家斯梅尔（Stephen Smale）为21世纪提出了18个问题（其中也有希尔伯特的老问题）；1998年，Schlumberger—Doll研究中心成立50周年时，邀请了一些科学家来会谈未来50年的科学与社会。他们报告的论题是：暴胀宇宙、太阳能、基因工程和互联网、富勒烯（C—60）材料、公共研究与工业、基因组与社会、全球变暖。我们这里看到的，是25个科学家谈“21世纪前50年的科学”（原书的副标题）。所谈的问题尽管也有人类永远好奇的宇宙和数学的问题，但更多的是关于我们自己的，如生命、意识、心理、道德、幸福、智慧、学习、遗传、医药、疾病……这一群第三文化的精英似乎要把我们对未来的幻想从天空拉回我们的生活。

的确，尽管科学在专业水平上离开我们越来越远，但它产生的结果和影响却离我们越来越近。远离我们的，会慢慢改变我们对自我和宇宙的认识（50年当然太短了）；走近我们的，每天都在改变我们每个人的生活和学习。正如科普作家T.Ferris在《谈科学普及》的演讲里说的，“科学还年轻。怀特海（Alfred North Whitehead）估计，一个新的思想模式渗透进一个文化的核心，需要1000年，而科学成为人们持续关心的事业，才不过500年。然而科学至少已经通过技术的、知识的和政治的三条途径，给世界带来了许多改变。”科学需要每一个人来感觉和加入那些改变。

25个科学家在这里想让大家明白的，就是那些改变；也许是“新文化”的特征，他们的谈话总的说来是小心翼翼的，没有多少浪漫的特别大胆的猜想，所预言的东西今天几乎都能看到影子。不过，我们还是可以听到一些令大脑兴奋的东西，如数学的量子化；基因活动的芯片；通过光经过空间的变化来检验引力的量子理论；重建侏罗纪公园；流动的信息束；超越我们的大脑的机器……而最惊人的，也许是Rodney Brooks的一句话，甚至可能超过当年达尔文宣布人是猴子的后代：“我们把自己作为一个物种的观念将发生改变；我们会发现自己不过是生产系统的一个部分而已。”

当然，同样的问题，不同科学家会有不同的预言。例如，Smolin在《宇宙的未来》里只提出了几个久远的问题，而不像Thorne在《时空未

来》那样浪漫地为即将来临的未来预写一章编年史；他预言量子计算将在未来50年成为现实，而有人认为未来量子理论的形式和基础需要更多“经典”方式的研究。Dawkins的“小摩尔定律”预言DNA排序速度会以指数形式增长，而更乐观的人认为它可能是“超指数”的——在2050年以前，会出现上万的转基因生物，我们的花园里会出现几千种陌生的花草树木。一些悲观的论调也可能令读者提不起精神来。例如，心理学家Paul Bloom说，“也许道德思想或意识的本性根本就超出了我们的理解力，不是因为它们处在什么特殊神秘的地位，而是因为我们没有理解那些事物的能力。我们大概像在努力理解微积分的狗。”另外还有很多“热的”问题没有谈到，例如大地、海洋、沙漠，市场、经济、战争。就是说，他们谈了很多自我，但几乎没有谈自我生存的环境。看来，还需要更多的科学家走上第三文化的舞台。

读者当然也会有自己的问题，而且应该提出自己的问题，这也是未来需要我们做的。（本来这是希尔伯特和爱因斯坦等前辈告诉我们的，可惜我们的学校没有听老人家们的话。）Roger C.Schank重新提出了教育的问题。在他看来，“新生的虚拟学校将取代现在的那些学校”“50年后的人可能会笑话我们今天还有老师、课堂和书本。”“教育将意味着，从两岁开始就在智慧——能回答问题并能提出新问题——指导下探索有意义的世界。一个个新的天地将向着好奇的孩子敞开。在这样的社会，教育关心的是你走进了什么样的虚拟（后来成为现实）的世界，在那些世界学会了多少事情。”他幻想的学习场景，大概是如今在校的同学（大同学和小同学）所向往的（至少比考试的课堂更讨人喜欢）。

本书最给人留下深刻印象的大概是未来的生活方式，而我们也许已经有了一点经历。翻译最后一篇《征服疾病》的时候，SARS来了，似乎在挑战作者的信心：“尽管我们现在担心会突然出现什么新的可怕的席卷全球的疾病，我们还是有理由相信，在未来的半个世纪，像艾滋病那样的毁灭性疾病的恐慌，也许没有，也许有一个——很可能没有。”我们不知道“也许有”哪一个，SARS似乎还不够可怕，但这个新世纪的过客真的为我们的生活带来了新世纪的些许色彩，让我们多少体验了未来生活的滋味。有的人因为恐惧而抑郁；有的人躲在小屋里，通过电话或电脑买东西；学生对着电脑屏幕上课……在大洋的彼岸，5个从中国实习回去的大学生被隔离起来。一个同学说，“我真想躺在椅子上看8个小时的电视——可要是72小时都那样，就一点儿也不好玩儿了。”如果说过去替科学担心的是战争与和平的选择，那么未来替科学

担心的可能是虚拟与现实的平衡。

这是25个人写的书，涉及很多的学科，本该由那些学科的专家来翻译，但我们不可能找那么多的人。幸运的是，我们现在的学习生活已经在“第三文化”的氛围里面，译者可以一边学习一边翻译。像书里讲的，“知识从四面墙壁涌来……是很容易获得的，只要你能大声地说出想知道的东西，四面的墙壁就能有回应”——虽然今天的四壁还没有隐藏那么多的宝藏，但是我们有专门的著作、期刊、百科全书，还有Internet和各种大众传播媒介。另外，如引言说的，“在这本书里，科学家并不是为了迎合大众才写得普及的，他们那样写，是为了吸引我们时代的论战中的其他学科的同行们。他们的目标不是科学的普及，而是一种努力的尝试，为的是不仅让广大的读者理解最新的科学研究，而且在科学本身的意义下把它说得通俗易懂。”因此，译者觉得有义务在传达文字之外还传播一点文化背景。那些背景的东西是“常识”，专家们当然不能在沙龙漫谈里罗嗦；也因为是常识，它们的来源也很多。学习和寻找来源的过程，可能也是未来学习生活的一个主要经历。虽然“寻章摘句”、“引经据典”是“传统的”学习，是未来的学习者可能嘲笑的对象，但是从读书的角度说，它似乎还有一定的需要。所以，译者在添加每一点注释的时候，都尽可能参考了不同来源的资料，并且尽可能利用原始的材料，所以注释里出现了许多引号；但是，一般情况下没有把出处写出来，以免离题太远；有时自己也发挥一点，不一定完全符合原来的某家之言；有时候指明了参考文献，包括专业论文和普及读物，供有兴趣的读者参考。我们能共享这些东西，要感谢所有传播它们的专家，其实他们早就走进第三文化的天地了。

[1] μ子是在宇宙线里发现的，除了质量大200倍，其他性质跟电子是一样的。它不是理论需要和预言的，所以理论家I.I.Rabi（诺贝尔物理学奖获得者）问，“谁让它来的？”——本书所有的注释都是译者添加的，以下不再说明。

[2] 我们知道，每个量子态都是若干（甚至无穷多个）状态的叠加，换句话说，它在同一时刻可以处在不同的状态。我们可以想象，假如计算机能同时处理多个状态，进行多个计算，它的威力该比0-1态的传统计算机大多少啊！2001年6月，一个丹麦科学家小组宣布他们发现了宏观气体的缠绕（entanglement，或译牵连、缠结，似乎还没有习惯说法）（Julsgaard et al.，Nature，27 September 2001），这是量子力学的奇妙现象：测量一个物体可以同时影响另一个与之没有牵连的独立物体。

[3] 最近有一些关于暗物质的发现，例如，在类星体周围发现了暗物质晕；暗物质可能聚集在我们星系的周围，像普通的恒星、行星和星际尘埃那样运动；凡是物质聚集的地方，暗物质的质量似乎总是物质质量的7倍（至少在星系的尺度上是这样的）。关于暗物质，目前还面临的一个挑战是区别不同的类型，如点状的neutralino（超对称理论预言的中性粒子，比质子重100—500倍）和延展的Q-ball（超夸克或超轻子的集合）。

[4] 物理学的无穷大有两个来源：点粒子假定和微扰论方法。弦理论将点粒子推广为可以延展的多维物质基元（如一维的弦和二维甚至更高维的膜），而圈量子引力是一个非微扰的理论。

[5] 《通向量子引力的三条途径》，李新洲，翟向华，刘道军译，上海科学技术出版社，2003年3月。

[6] SETI就是Search for Extraterrestrial Intelligence（寻找地外智能）。这个计划从20世纪60年代就开始了。

[7] 古斯说，大爆炸理论并不是真正关于宇宙爆炸的理论，而是关于爆炸以后的宇宙的理论。它不能说明为什么爆炸，爆炸以前发生了什么。暴胀理论有可能回答宇宙膨胀的动力是什么，还可能回答宇宙的物质是从哪里来的。50年后，“我们可能知道哈勃常数的大小，知道宇宙的形状是开放的还是封闭的，知道宇宙的膨胀是否发生过变化，但是，我们也许还是不知道宇宙真正的起源。”

[8] William Olaf Stapledon（1886—1950）是英国科幻小说家，他的许多幻想成为后来科学家的思想来源；这里引的《造星者》（Star Maker（1937））还启发Freeman Dyson提出了“戴森球”（Dyson sphere）的概念：一个人造的包围恒星的球壳，几乎能吸纳所有的辐射能量，人类生活在这样的球壳上才可能有足够的能源。

[9] 这不是说我们要拒绝“好的科学假说”，而接受什么“坏的”。它的意思是，一个科学的假说应该具有能被否定的“素质”，否则它就不是好的科学假说（甚至是伪科学的）——这是著名的波普尔（K.Poper）的“证伪”观。

[10] 行星轨道也许真的存在什么法则。人们在18世纪发现，如果用天文单位（日地距离）来计算，行星轨道大小之间存在这样的联系：$\alpha_n = 0.4+0.3\times 2\ n^{-2}$，其中n为行星序号，不过水星序号为-∞。这就是有名的Titius-Bode法则。后来，在n=5的地方发现了小行星带；在n=8的地方发现了天王星；在约相当于n=9的地方发现了海王星。但是，冥王星却不在n=10的地方（已经有人认为它不是真正的行星）。太阳系的起源理论应该为这个法则找一个根据。

[11] 1877年，斯基帕雷利（Giovanni Schiaparelli，1835—1910）在火星大冲时——也就是地球与火星处在太阳同一侧面且在同一条直线上（幸运的是，2003年8月27日是近6 000年来火星最接近地球的时间，被定为“世界火星日”。实际上，在整个2003年秋季，都能在南方的天空看到又大又亮的火星）——发现其表面有许多天然河道，意大利语是canali，被英国报纸错译为canal，成了大家知道的“运河”（或“沟渠”）。相信火星存在生命的名人之一，就是后面的篇章中提到的美国天文学家洛维尔（Percival Lowell）。

[12] Simon Singh，*Fermat’s Last Theorem, The story of a riddle that confounded the world’s great minds for358years*，Fourth Estate Lt.，London，1997.（中译本《费马大定理》，薛密译，上海译文出版社，1998）很好讲述了怀尔斯和费马大定理的故事；怀尔斯的两篇论文发表在*Annals of Mathematics* 142（1995）：443—551，553—572。关于怀尔斯的证明，至今还能听到怀疑的声音。

[13] 为了迎接新千年的到来，克莱数学研究所悬赏了7个数学难题。下面根据克莱研究所网站提供的材料简单介绍一下这几个问题：（1）我们都熟悉勾股定理，$x^2+y^2=z^2$，关于它的整数解的问题在欧几里得时代就解决了。但是，类似的更复杂的方程却没有一般的解决办法（希尔伯特第10个问题）。伯奇/斯温纳顿-代尔猜想，假如方程的解是阿贝尔型的变量，则解的数量跟ζ函数在s=1附近的行为有关[所谓ζ函数即ζ（s）$=1/2^s+1/3^s+1/4^s+\ldots\ldots$s是复数]。假如ζ（1）=0，有理数解有无穷多；否则，解的数目是有限的。（2）20世纪的数学家发现，可以把简单的几何碎片粘结起来逼近某个给定的几何体，但有时粘结的碎片却没有几何意义。霍吉猜想说，对某些“好的”空间形式（投影代数空间），那种被称作“霍吉圆”的碎片实际上是所谓代数圆的线性组合。（3）尽管作为流体力学基础的纳维尔-斯托克斯方程已经写下100多年了，我们对它的了解依然浅薄，我们特别希望能从这个方程的数学理论认识湍流。（4）有些问题的答案检验起来很容易，但计算机做起来却需要几乎无限的时间，这就是所谓的NP问题。P问题则可以通过“多项式”的时间算法来计算。这里，P=Polynomial（多项式），NP=Nondeterministic—Polynomial（非确定多项式）。（有意思的是，据说在克莱悬赏之前，中国学者已经把这个问题解决了。）（5）我们知道二维球面（如地球表面）是单连通的（可以收缩为一个点），庞加莱在100年前问，三维球面是怎样的情形呢？（6）黎曼猜想说，ζ函数的非平凡零点都落在实部为1/2的一条直线上。这个猜想联系着许多关于素数分布的难题，例如，哥德巴赫猜想不过是它的一个特例。（7）用杨振宁—米尔斯（Mills）的规范场理论来描写基本粒子的强相互作用时，需要一种微妙的量子性质，即所谓的“质量间隙”：尽管经典的波动以光速运动（质量为0），然而量子粒子却具有正的质量。我们在理论上还不能理解这一点。

[14]　卢瑟福嘲笑物理学以外的事情都是“集邮”，而生物学家认为他们的集邮活动是认识生物多样性所必需的，而且永远不会到头。在《终极理论之梦》中，温伯格为了超级对撞机也攻击了生物分类的研究。数学家在这里把两家都批判了。今天科学家们在科学外的这些争吵（主要是为了争取经费），大概会成为未来科学史或科学学的有趣话题。

[15]　30年战争，是1618—1648年间为争夺欧洲霸权而发生的一场全欧国际性大混战。它结束了自中世纪以来“一个教皇、一个皇帝”统治欧洲的局面，德国分为近300个独立的大小不同的诸侯领地和100多个独立的骑士领土，皇帝在欧洲恢复天主教地位的企图完全破灭，神圣罗马帝国事实上已不复存在。“在整整一个世纪里，德意志被历史上空前未有的最无纪律的暴兵纵横反复地蹂躏着……物质的破坏，人口的凋零，是无穷无尽的。当和平到来的时候，德国已经不可救药了，已经被踏碎、被撕破、遍身流血，躺倒地下了……”（恩格斯）。

[16]　“天球的乐音”和“爱的力量”是西方现代科学诞生之前人们对宇宙的认识。毕达哥拉斯的“天球的乐音”反映的是宇宙的和谐；“爱的力量”则是从太阳的崇拜产生的。例如，文艺复兴时期的菲西诺（Marsilio　Ficino，1433—1499）在太阳叙事诗中写道：太阳“像抚爱一样温柔而不知不觉地渗出万物……因而是宇宙的创造者和推动者……同样，上帝自己也是无处不在的，它也抚育和滋养万物。它不是靠强力来创造和推动万物，而是靠伴随着它的爱。”这实际上也是布鲁诺的思想根源。

[17]　　Gaia是希腊神话里的大地女神，她从卡俄斯（chaos，创世之前的空间，也是我们现在所谓“混沌”一词的来源）产生，又产生了天空、大地和海洋。Tellus是同一个女神在罗马神话里的名字。

[18] “外边的黑暗”，见《新约·马太福音》（22：13）：“捆起他的手脚来，把他丢在外边的黑暗里”（多数人认为“外边的黑暗”指地狱，也有人提出它是“天堂的郊外”）。

[19] Navajo是美国的一个印第安部落。

[20]　Douglas　Adams（1952年3月11日—2001年5月11日）是当代西方影响巨大的幽默讽刺剧作家，他的喜剧不是发生在现实世界，而是在科学幻想的舞台。他80年代的《银河顺风车旅行指南》（*The Hitchhiker’s Guide to the Galaxy*）系列今天依然是令人难忘的经典。

[21] 这是第20章的一个情节：海蒂在树丛中摘栗子，亚当来了。“奇怪的是她竟然没有听见他来了！也许是树叶的沙沙声太响。当她发觉有人走近，吓了一跳——装着栗子的盆子都吓掉了；等她发现是亚当时，灰白的脸一下子变红了。那羞怯……”

[22] Hugh Lofting（1886—1947）生在伦敦，16岁到麻省理工学院读书，没有学完就回国了。后来做过土木工程师。杜里特（Dolittle，意思大概是“无为”）医生的系列故事是他在第一次世界大战期间为了与两个儿子沟通而创作的。近年，美国20世纪福克斯公司为这位大名鼎鼎的医生新编了两部电影，他可以用428种动物的语言与动物交谈。

[23]　“道奇”鼠（Doogie）是钱卓博士领导的一个科学家小组（成员分别来自普林斯顿大学、麻

省理工学院和华盛顿大学）通过基因改造而“生产”的系列实验小老鼠。如今关在普林斯顿大学生物系的笼子里。名字仿照美国电视剧Doogie Howser，主人公是位天才儿童，二十来岁就当了住院医生。不过，从聪明小鼠到天才儿童的路还不知道有多长。正如钱博士说的，“Doogie在实验室所表现的优势，是具有较强的觅食与逃离危险的能力，但仍然只是老鼠程度的聪明，绝不会因此就变成钢琴家。”他们的文章发表在1999年9月1日的《自然》杂志上：Ya—Ping Tang, Eiji Shimizu, Gilles R.Dube, et al.，Genetic enhancement of learning and memory in mice.*Nature* 401，63—69（1999）.

[24] 阿尔茨海默症是德国精神病理学家Alois Alzheimer在1907年首先发现的一种老年痴呆症，多发生在65岁以上的老人（女性患者更多）。

[25] 希腊神话故事：忒修斯怀着悲痛埋葬了父亲，然后将阿提喀的童男童女乘坐的那只船献给阿波罗，那是一只能容纳30个水手的船。雅典人为怀念这次神奇的历险，设法保全这只船，把船上的朽木不断地更换。因此，许多年以后，在亚历山大大帝时还可以看到这一古老而珍贵的纪念物。

[26] 波兰斯基（Roman Polanski）是美国当代大导演，《房客》（*The Tenant* ）是他70年代的作品。他的《钢琴家》获2002年度美国电影评论家协会四项大奖和第75届奥斯卡最佳导演奖。

[27] 盖奇原是新英格兰维蒙特地区的一个铁路工头，1848年9月，因为意外的爆炸，一根铁杆从他左颊穿过前额。令人惊奇的是，他竟活了下来，而且两个月就恢复了体格。可是，性格却完全改变了，“强壮的身体和野兽的性情”。120年后，蒂梅修博士（Damasio，不是正文说的那一个）和她的同事重组了盖奇头颅的三维图像。现在一般认为，大脑里有一个专门控制情感的区域，不会影响其他功能。

[28] Miracle—Gro，一种很出名的水溶性植物肥料。Gro大概是grow（生长）的缩写，不知道园艺界是如何翻译这个品牌的。

[29] 爱因斯坦这句有名的原话是“Raffiniert ist der Herrgott aber boshaft ist er nicht”。现在还镌刻在普林斯顿大学数学系费因楼（Fin Hall）大厅的壁炉上。

[30] 贝叶斯网络图（Bayes net）是以贝叶斯概率为基础的图形化的推理网络。

[31] 1950年英国数学家图灵（Alan Mathison Turing，1912—1954）在《计算机械与智能》一文中提出了一个“试验”（Turing test），让一个人跟一台机器和另一个人交流，假如他分不清机器与人，就说那样的机器具有与人一样的思维能力。

[32] Burrhus Frederick Skinner（1904—1990）继承和发展了华生（J.B.Watson）的“刺激-反应”的公式，把它发展为新行为主义，认为动物和人类的学习是一种操作，还把教学程序归结为“刺激-反应-强化”，认为在教学中采取积极强化措施可以使某种行为形成习惯。他喜欢举饿老鼠获得食物的例子来说明他的行为主义观点。

[33] 洛克（Locke），这里应该说的是哲学家John Locke（1632—1704），认为心灵是一块白板，后天的经验才是认识的源泉。哲学方面有著名的《人类理解论》。

[34] William James（1842—1909）是美国哲学家和心理学家，他的《心理学原理》是美国人自己的第一本有完整体系的心理学著作。“意识流”的概念就第一次出现在这部书里。他自己说，“记忆、思维和情感存在于基本意识之外，这一发现乃是我从事心理学研究以来这门科学所取得的最重大的进展。”关于James和他那出名的家族（父亲是神学家，弟弟Henry是大作家，下一篇文章就会提到他），近年出版了一本新传记，H.M.Feinstein, Being William James（《就这样，成了威廉·詹姆斯》，季广茂译，北京：东方出版社，2001年2月）。

[35] 乔姆斯基（Noam Chomsky）在20世纪50年代提出转换生成语法体系，在哲学、心理学、逻辑学、语言教学、文体理论等多个领域都有着深远影响，今天还成为复杂系统研究中形式语言的基础。他还是积极的社会评论家，被认为是“美国最伟大的异议分子”和“美国人的良心”。最近Robert F.Barsky为他写的传记就是《乔姆斯基：不驯的生命》（Noam Chomsky：A Life of Dissent）。

[36] Wilson有时被称为“社会生物学之父”。他主张，只有承认所有问题都能溯源到科学，我们才可能认识周围的世界。他认为，战争和侵略是人的天性，都源于“适者生存”的老观念。关于“利他”，他说“同情是有选择的，而且终归是为了自己。”他遭人反对，还因为他相信男人与女人扮演着根本不同的生物角色；宗教信仰都可以归结为进化论原理。

[37] 《芝麻街》（*Sesame Street*）是美国公共广播公司（PBS）著名的儿童益智节目，1969年在纽约胡伯斯问世，目前已在全世界140多个国家播映，儿童观众已超过1.3亿。《罗杰斯先生的邻居》是PBS 1968年推出的儿童节目。

[38] 存在性恐惧（existential dread）指伴随人的生存而存在的那些无端的恐惧和焦虑。这个概念源于存在主义哲学的先驱、丹麦哲学家克尔凯郭尔（Soren Kierkegaard，1813—1855），他认为“存在”就是由痛苦、烦恼、孤独、绝望等情绪构成的个人的存在。

[39] 亨利·詹姆斯（Henry James，1843—1916）开创了心理分析小说的先河，在他的笔下，出现了仿佛是迷宫般的普通人的内心世界。在《鸽翼》（The Wings of the Dove，1902）中，他发掘了人物“最幽微、最朦胧的”思想和感觉，把“太空中跳动的脉搏”转化为形象。在兰登书屋（Random House）1996年评选的20世纪百部最佳英文小说中，亨利个人就占了三部。

[40] Viacom是全球第三大媒体和娱乐公司，著名的CBS（哥伦比亚广播公司）就在它的旗下。

[41] 法国大数学家阿达玛（J.Hadamard）对“内省”（introspection）有过简明的解释：内省的心理学方法就是主观的方法，即“从内部观察”的方法。“有关思维方式的信息直接来自思想者本人——他通过观察内心来报告自己的心理过程。”

[42] 华生（John Broadus Watson，1878—1958）是行为主义心理学创始人，他根据对动物和儿童心理的研究，主张作为自然科学的心理学应该研究行为，而不应该研究意识。

[43] 帕索里尼（Pier Paolo Pasolini，1922—1975）是意大利作家，小说重视视觉映像，成了电影导演喜欢的题材。后来他自己也做了导演。作品有《一千零一夜》《马太福音》《十日谈》等。1975年被同性恋少年杀害。

[44] 20世纪初，英国统计学家斯皮尔曼（Charles Spearman，因子分析法的发明者）发现，不同的智力测试的结果中存在一个共同因子，他称之为g，代表一般性智力。最近的研究似乎证明智力是大脑特殊区域产生的，可能真的存在那样一个一般性的因子。

[45] 兰德（Ayn Rand，1905—1982）的哲学是“人为自己而活”，人的最高目标在于追求“理性的自利”。她主张理智、自我、个人主义，极力推崇“自由放任”。她的哲学为某些贪婪的野心带来了一丝道德的安慰。

[46] “小人理论”认为在我们大脑的某处有一个小矮人，试图模仿大脑正在进行的活动。克里克（Francis Crick）在《惊人的假说》（见《第一推动》丛书）中称它为“小矮人谬误”（the Fallacy of the Homunculus）。这是一个诱人的伪科学的理论，与神创论、活力论、外星人杂交论等没有本质的区别，它们都是以一个更大的问题来替代我们的传统问题。

[47] “流”或“畅”（flow）是西氏提出的一个概念：“具有适当的挑战性而能让一个人深深沉浸于其中，以至忘记了时间的流逝、意识不到自己的存在的体验。”这里提到的两本关于“流”的书，在我国台湾出版的中译名分别是：《快乐：从心开始》《生命的心流》，很传神。

[48] 1929年，哈佛大学的工程师德林克（Phillip Drinker）用旧真空吸尘器的零件做了一个呼吸器，通过真空管有节奏地把空气吸入和排出病人体内。在索尔克发现脊髓灰质炎的疫苗之前，它是脊髓灰质炎患者的唯一希望。

[49] 1952年3月26日，匹兹堡大学的索尔克博士培育的一种预防脊髓灰质炎的疫苗在90个成人和儿童身上试验成功。后来，辛辛那提大学的萨宾研制了口服用脊髓灰质炎减毒活疫苗，也就是今天的小儿麻痹症口服糖丸。然而，1954年诺贝尔生理学或医学奖授予了首次在非神经组织中培养出脊髓灰质炎病毒的三个科学家，却没有给他们两位。

[50] Major depression是《美国精神病协会精神疾病诊断和统计手册》（DSM—III，1980）首先使用的，在中文里似乎还没有约定的名词（有“重性抑郁症”“主要抑郁症”等说法）。译者觉得这里应该用一个更简单响亮的名字，当然它不符合医生的规矩。

[51] 1940年，纳粹德国疯狂迫害犹太人，美国政府不想因为犹太难民问题与希特勒和傀儡的维希法国闹僵关系，要求各地使领馆缩减大多数赴美犹太人配额。刚上任的驻马赛领事海勒姆·宾厄姆（Hiram Bingham）却“违反了美国的反难民政策”，在一年之内救助了2500多名犹太人（包括著名艺术家马克·夏加尔、马塞尔·杜尚、马克斯·厄恩斯特等）。60年后，美国政府表彰了他们的这位辛德勒式的英雄。

[52] 在美国，过去在招工广告里常常用大字写着“不招爱尔兰人”。

[53]　直到今天，在印度一些地方还遗存着这样的风俗：妻子在火葬丈夫时自愿投身烈火，这样可以赢得“荣誉和尊重”，据说还可能阻止灾难的发生。

[54] 杜奇曼夫人是历史学家，她的《八月的炮声》（*Guns in the August* ）曾获美国普里策奖。在《遥远的镜子：多难的14世纪》（*A Distant Mirror*：*The Calamitous 14 th Century* ）里，她以14世纪的灾难岁月来映照20世纪的人类社会。

[55] 我们还记得这些事情：科索沃战争、美国校园的枪声、克林顿的性丑闻……

[56]　1979年10月26日，世界卫生组织宣布：人类最后一个天花病人，来自“非洲之角”索马里的牧民——阿里·毛·马林，在1977年被治愈了。1980年5月，联合国在内罗毕宣告：“天花已经在世界上绝迹。”

[57]　马基雅维里（NicoIo　Machiavelli，1469—1527）是意大利政治家和思想家。在《君主论》（潘汉典译，商务印书馆，1985）中，他提出为了目的可以不择手段的著名思想。他还有一个有趣的比喻说，“君主必须是一头狐狸以便认识陷阱，同时又必须是一头狮子，以便使豺狼惊骇。”

[58]　费米与原子弹的故事可以参见他夫人Laura　Fermi写的*Atoms in the Family*：*My Life with Enrico Fermi* （《原子在我家中》，何芬奇译，科学出版社，1979；《费米传》，何兆武、何芬奇译，商务印书馆，1997）。

[59] 借萨根（Karl Sagan）的话说，在古今学者中像布罗诺夫斯基那样博学的还为数不多。他认为人类一切知识都是有趣的而且可以学会的。他的《人之上升》（*The Ascent of Man* ）出版以后（四川人民出版社20世纪80年代出版过中文本；后来海南出版社出版的中译本为《科学进化史》），被改编为电视节目，也同萨根的《宇宙》一样反响强烈。

[60] 原文“Nonlinearity giveth chaos, and nonlinearity taketh it away.”很有趣，显然借了《圣经》的句式，如“the Lord giveth, and the Lord taketh it away.”（Job 1：21）。

[61] “6度分离”（six degrees of separation）是哈佛大学的米尔格拉姆（1933—1984，读者应该记得，前面讲过他有名的“服从权威”的实验）在1967年提出的。大概意思说，地球上任意两个陌生人平均只需经过6步就能发生联系。照他的说法，世界真是太小了。近几年来，本文作者和他的学生Duncan Watts对这个“小世界效应”进行了模型研究。

[62] 这是费曼在他著名的《物理学讲义》第二卷的结束语。

[63]　在东南亚的边远地方，生存着一类萤火虫，成千上万的它们能以相同的节律闪光，这种行为大不同于自然界的其他同步现象。在1993年12月的《科学美国人》上，作者有一篇谈生物同步现象的文章：《耦合振子与生物学同步》Coupled Oscillators and Biological Synchronization.Scientific American.1993（12）：68—75）。

[64] 薛定谔的《生命是什么》新译本已经在我们的《第一推动》丛书里出版了。

[65] 维特根斯坦一生提出过两个完全不同的哲学。以《逻辑哲学论》（1922）为代表的前期思想认为，语言与事实有着一一对应的关系，语言是“反映”。后来，在《哲学研究》（1953）中，他认为语言是“活动”（或“游戏”），词和句子在不同的环境有不同的用法。于是，哲学的任务不是分析语言的意义，而是描述语言的用法（“语言游戏”的德文原词为Sprachspiel）。

[66] 简单说来，反应坐标就是势能曲面所决定的能量最低的反应路径。

[67] 此书中译本在我们的《第一推动丛书》中已经出版过（李绍明、徐彬译）。

[68] 梅达瓦（Peter B.Medawar，1915—1987）是英国动物学家，因为发现胚胎期的获得性免疫和免疫耐受性，“为实验生物学开辟了一条新路”，与Sir Frank Macfarlane Burnet分享了1960年诺贝尔生理学或医学奖。他的《溶解术》（The Art of the Soluble）同沃森的《双螺旋》一起，被蓝登书屋选为20世纪百部最佳非小说类图书。

[69] Watson J.and Crick F.，A Structure for Deoxyribose Nucleic Acid.Nature 171，737（1953）.最近，《自然》杂志专门纪念了这篇文章发表50周年。

[70] 那小虫叫秀丽隐杆线虫，非常适合真核生物进化的遗传学、分子学和细胞学的研究。

[71] 作者原来说的是“摩尔定律的儿子”（Son of Moore’s Law）。1965年，Intel公司创始人之一的Gordon Moore讨论集成电路的未来时，预言在未来10年里，集成电路的组成每年将增加1倍。后来，预言成为计算机科学中有名的Moore定律。在过去的30多年里，这个经验法则惊人地准确，不过翻倍时间逐渐延长到18—24个月。现在我们还有一个关于互联网的Metcalfe定律：网的价值与上网人数的平方成正比。

[72] Queen Elizabeth II是美国卡那海运（Cunard Line）拥有的世界唯一一艘横渡大西洋的豪华游轮，也是“全世界最有名的游轮”。

[73] Evans（1931—1979），《微电脑世纪》（*The Micro Millennium* ）。

[74] Sydney Brenner 1927年出生在南非，1954年在牛津获博士学位。1996年他建立了加州伯克利分子科学研究所。2002年10月7日，他终于和另外两个同样以研究线虫出名的科学家（H.Robert Horvit和John E.Solstun），“因为在器官发育及程序性细胞死亡的基因调控方面的贡献”获得了诺贝尔生理学或医学奖。

[75] 本书的副标题说明了它的内容：“科学、错觉和好奇心”（*Science, Delusion, and the Appetite of Wonder* ）。有人评论这可以当作萨根《魔鬼出没的世界》（*The Demon-Hunted World* ）的续篇。

[76] DNA是一连串核苷酸按顺序连接起来的有限长度的分子链，因此有开始，也有终止。在64个密码子中，有3个没有对应的氨基酸，最初被认为是“无意义密码子”。后来发现它们代表链的

终止（蛋白质的合成到此告一段落）。在基因突变中，如果决定某一氨基酸的密码子变成一个终止密码子，我们就说那是“无义突变”；密码子变为UAG的无义突变就是琥珀突变。

[77] Sydney Brenner，“Theoretical Biology in the Third Millennium”，*PhilosophicalTransactionsof Royal Society*，B，1999.

[78] 属于人科动物的南方古猿生活在上新世到更新世，化石几乎都是在非洲发现的，距今约50万—370万年。1974年，美国D.C.Johnson在埃塞俄比亚阿法地区发现了一具成年女性南方古猿的化石骨架，发现的那天晚上，发掘队员们高兴地播放着甲壳虫乐队的流行歌曲“天上带着宝石的露茜”，于是有了那骨架的绰号。露茜骨架完整，年代久远（距今约350万年），可以肯定是最早的人类祖先。遗憾的是，露茜的头骨前部完全缺失了，复原的头跟垒球大小差不多，很难测定脑量。

[79] *The War of the Worlds*，描写的是高度发达的火星人入侵地球，打碎了地球上的社会制度。

[80] 贝格尔1号当然是达尔文1831年环球远航为进化论找证据的那艘军舰（Beagle在英文里是一种小猎犬）。贝格尔2号由英国莱斯特大学、开放大学等高校共同设计，肩负着更大的使命。2003年6月3日，它搭欧洲航天局的“火星特快”飞船，在哈萨克斯坦的拜科努尔发射场由俄罗斯的联盟号火箭送上天了。

[81] ALH（Allan Hills）84001是美国的一个陨石搜寻小组1984年在南极洲发现的，重1.93千克，说它来自火星的最大证据是它包含着接近火星大气的气体。

[82] 这是John Templeton爵士在1972年创立的一个“宗教进步奖”（Templeton Prize for Progress in Religion），也被誉为宗教的诺贝尔奖。获奖者除了宗教界人物，也有不少是在科学与宗教的对话方面做出过重大贡献的科学家，特别是物理学家。

[83] 克拉克（1917—）是英国科幻小说大家，他的书发行了5000万册，最著名的是《太空漫游》系列。据说，从他头发取出的DNA即将真的“漫游太空”。

[84] 凡尔纳这本书当年没能出版，手稿也失踪了，1994年才找到。小说是反乌托邦的，预言了传真机、电子计算机、汽车、地铁等。但那个巴黎是技术的世界，没有一点儿艺术。

[85] Amazon.com是世界最大的网上书店。（书店的创业者们当初实在想不出一个好名字，就用了“亚马逊”，没有什么特别的意思。）e—Bay是美国人Pierre Omidyar于1995年5月1日创办的一个拍卖网站，是目前世界上注册人数最多（超过1000万）的拍卖网站。

[86] 1956年，人工智能的先驱者西蒙（Herbert Simon）、明斯基（Marvin Minsky）、麦卡锡（John McCarthy）等人在达特茅斯学院（Dartmouth College）召开了第一次人工智能研讨会，“人工智能”一词也在会上第一次出现，拉开了AI研究的序幕。

[87] 1947年10月14日，美国贝尔公司在加州试飞了第一架超音速飞机，它的名字叫X-1，不过，

它是被挂在一架巨大的B-29轰炸机上飞上天的。这个系列一直延续着，因为都是试验性的飞机（所以才叫X），所以常常被“抛弃”。“黑鸟”（blackbird）是美军高速侦察机。

[88] 这原是一句英国谚语，“An Englishman’s home is his castle”，照英国的法律和习俗，个人的家是不得随便闯入的。

[89] 《1984》是英国小说家George Orwell（1903—1950）1949年写的政治幻想小说，描写了一个反乌托邦式的极权社会，那个社会流行的口号是，“战争是和平，自由是奴隶，无知是力量”，个人的权利被完全践踏了。

[90] 汪达尔人（Vandal）原是日耳曼民族的一支，公元455年他们占领罗马，毁坏了大量公共建筑，“悲惨地看过古罗马废墟的游客很容易情不自禁地指责古代的哥特人和汪达尔人”（E.Gibbon《罗马帝国衰亡史》）。由此派生的Vandalism指狂热的革命者对艺术的破坏。

[91] 富勒烯（fullerenes）是1985年由Curl, Kroto, Smalley偶然发现的（因此获1996年诺贝尔化学奖），因为具有完美的球形（竹笼）结构，跟建筑师Buckminster Fuller的短程线圆顶结构异曲同工，所以三个发现者就借Fuller的名字来命名它。它是当前国际化学、物理学和材料科学共同关心的热点。（富勒烯的故事，可以看J.E.Baggott, Perfect Symmetry：The Accidental Discovery of Buckminsterfullerene或《完美的对称》，李涛、曹志良译，《哲人石》丛书，上海科技教育出版社，1999。）

[92] 2003年的诺贝尔医学或生理学奖授予了Paul C.Lauterbur和Peter Mansfield，表彰他们“关于磁共振图像的发现”。

[93] 《西方世界经典》，*TheGreatBooksoftheWesternWorld* ，Encyclopaedia Britannica, Inc.，Chicago，London，Toronto，1952.丛书共54卷，前3卷是一般介绍，后51卷收录自荷马以来3 000年的70多位西方文化大师的著作。读者看看熟悉多少，又读过多少：Homer, Aeschylus, Sophocles, Euripides, Aristophanes, Herodotus, Thucydides, Plato, Aristotle, Hippocrates, Galen, Euclid, Archimedes, Apollonius, Nicomachus, Lucretius, Epictetus, Marcus Aurelius, Virgil, Plutarch, Tacitus, Ptolemy, Copernicus, Kepler, Plotinus, Augustine, Thomas Aquinas, Dante, Chaucer, Machiavelli, Hobbes, Rabelais, Montaigne, Shakespeare, Gilbert, Galileo, Harvey, Cervantes, Francis Bacon, Descartes, Spinoza, Milton, Pascal, Newton, Huygens, Locke, Berkeley, Hume, Swift, Sterne, Fielding, Montesquiei, Rousseau, Adam Smith, Gibbon, Kant, J.S.Mill, Boswell, Lavoisier, Fourier, Faraday, Hegel, Goethe, Melville, Darwin, Marx, Engels, Tolstoy, Dostoevsky, William James, Freud. 除了这些个人而外，还有一卷是美国的联邦党人文选。

[94] 塔利（Tully）是西塞罗（Marcus Tullius Cicero，106B.C.—43 B.C.）的英文名字。

[95] Jeopardy（杰帕迪）是美国一个很有名的智力竞赛电视节目，已经搞了很多年，来自全国各地的许多聪明人都会到这个节目中来较量。

[96] 尼尔（Alexander Sutherland Neil，1883—1972）是英国教育家，主张儿童是自由的，一切

纪律、道德的训练都应该放弃；杜威（1859—1952）是有名的实用主义哲学家，他主张教育即生活，应该“从做中学”。

[97] SAT=Scholastic Aptitude Test，美国的学术智能测试。

[98] 作者在1989年第一个提出了“虚拟现实”（Virtual Reality）的名词（尽管相关研究在20世纪60年代就有了）。关于VR，可以参考一本《虚拟现实的形而上学》：Michael Heim，*The Metaphysics of Virtual Reality*（New York：Oxford University Press，1993）

[99] 作者在另一篇文章里的话更清楚地说明了这一点：“丰满的软件的关键在于它开发利用了多少东西。假如摩尔定律在未来二三十年还成立，那么不仅在我们的行星地球会出现大量计算，计算的维护还将费尽几乎每个人的努力。这样，我们面对着一个‘帮助台的行星’。”——我们甚至可以将整个地球当作一个巨大的帮助台，地球上的每一个人都是需要帮助的。

[100] 混沌学的先驱者、气象学家洛伦兹在1963年向纽约科学院报告说，“海鸥拍打翅膀就可能永远改变气候的过程”。1972年12月在美国科学促进会（AAAS）讲话时，他的题目更有诗意了：“可预测性：一只蝴蝶在巴西拍动一下翅膀能否在得克萨斯引发龙卷风？”蝴蝶效应说的是系统演化对初始条件的敏感性，是我们最早发现的混沌特征。

[101] Causal potential大概指一个原因引发结果的潜在力量。这里试译作“因势”，可以与数学物理中的各种势函数发生联想，实际上它们都是一样的意思。

[102] Macintosh是美国苹果公司1984年推出的个人计算机，名称来自加拿大John McIntosh在1796年培植的一种红苹果。所以机器的名称实际上就是“红苹果”。

[103] Michael Behe在*Darwin’s Black Box-The Biochemical Challenge to Evolution* -Free Press，1996（《达尔文的黑匣子：生化理论对进化论的挑战》，刑锡范等译，中央编译出版社，1998）中以生命系统的“不可约减的复杂性（irreducible complexity）”来证明达尔文的进化论到头了。

[104] 博尔赫斯（1899—1986）生长在阿根廷，是“影响欧美的第一位拉丁美洲作家”，其作品近年也在中国流行。他说“天堂应该是图书馆的模样”。他在《通天塔图书馆》里描绘了一个“宇宙”，“别人管它叫图书馆，由许多六角形的回廊组成，数目不能确定，也许是无限的……”小说最后说，“图书馆是无限的，周而复始的。假如一个永恒的旅人从任何方向穿过去，几个世纪后他将发现同样的书会以同样的无序重复着，重复后便成了有序：宇宙的秩序”。

[105] 道金斯（Richard Dawkins）在《自私的基因》（卢允中、张岱云译，吉林人民出版社，1999）中提出，选择的基本层次是独立的各个基因，而作为个体的人，最好被理解为“生存机器”，或者是为了它们自己的繁衍而在一起运作的基因的聚合体。

[106] 遗留（legacy）指的是长期使用的计算机系统所建立起来的服务和功能，如果新的客户想以新的系统来取代它，费用会很高。遗留（或既有、保留）系统往往成为开发新系统的阻碍。

[107] 远程沉浸（tele-immersion，或译远距离兼容、远距实境等）就是让不同地方的电脑用户能同时在一个虚拟的环境下协同工作，就像沉浸在同一个生活空间。这个名词是伊利诺州大学芝加哥分校电子可视化实验室EVL（Electronic Visualization Laboratory）在1996年提出来的。

[108] Fred Astaire（1899—1987）是美国一代舞王，被称为“世界上最伟大的舞蹈家”。他的表演、歌唱和舞蹈改变了美国音乐电影的面貌。主要作品有*Lady Be Good*！*Flying Down to Rio*；*Shall We Dance?Hot Hat*；*Funny Face*；*Dancing Lady* 等。

[109] 古尔德（Glenn Gould，1932—1982）生在加拿大，1964年4月在洛杉矶演出后就告别了短短9年的舞台生涯，大概是音乐史上第一个拒绝舞台的演奏家。他是完美主义者，认为只有选用录音带上最好的部分（他常常亲自做剪辑）才能实现完美的音乐。他说，“技术解放了艺术家，使他有更多的时间和自由，以自己的最高水平构思一部作品，远离紧张、焦虑、手指错音等枝节，达到完美的境地；技术能改变音乐会上令人讨厌而又不可避免的不确定性，将与音乐无关的个人东西排除在音乐之外。”

[110] 美国C—SPAN是Brian P.Lamb在1979年创立的公共事务有线电视网（Cable Satellite Public Affairs Networks），又称国会台，大量如实报道美国政治生活的大事，不作任何编辑、评论和分析。Lamb自己说，“在电子媒介的历史上，只有C-SPAN能让一些人站在麦克风前不受干扰地表达他的思想。”

[111] 在莎士比亚《麦克白》第5幕第3场中，麦克白对太医说，“你难道不能治疗那病态的心理，从记忆里拔去生根的忧郁，抹去写在脑膜的烦恼，拿一剂甘甜的忘忧药，把那淤积在胸间、重压在心头的毒素清洗干净吗？”（本节原题smart drug，译者根据内容把它改成现在这个在汉语中似乎更好听的词儿。）

[112] denial原是心理学家A.A.Brill 1914年翻译弗洛伊德《日常生活的精神病理学》时用的一个术语，20世纪80年代，in denial却在美国流行开来。说的是一种无意识的抵制情绪，不愿承认痛苦的现实。

[113] 在下面我们会看到，哈里斯提出了一种父母对孩子成长的“零影响”假设；她认为过去的“教育假设”是偏见。

[114] Steven Pinker，*The Blank Slate*：*The Modern Denial of Human Nature.* 作者批判了婴儿大脑是“空白的石板”的观点，认为人类先天就有物种生存所需要的普遍结构。

[115] 1856年在德国Neander山谷发现第一个尼安德特人（大约在30万年前出现，3.5万年前灭绝）骨架后，人们都相信它是人类的祖先。1997年，德国科学家发现它的DNA与人类的不同，而是另外一种灭绝了的动物。

[116] 1998年，Rhone—Poulenc与德国Hoechst合并为Aventis，方向是医药、疫苗和生物工程。

[117] 这是目前发现的在突变时引起家族性阿尔茨海默症的三个基因：β-淀粉样蛋白前体

（APP）、衰老前素-1（PS1）和衰老前素-2（PS2）。

[118] 哈西德派（Hasidism）是犹太教的一个派别，重视情感价值和虔诚，也主张快乐和积极的爱，奥修说它是“犹太教的精华”。Martin Buber（1878—1965）生在维也纳，犹太哲学家、神学家和社会活动家，20多岁就成了哈西德派。他最有名的著作是关于“对话哲学”的《我和你》（*I and Thou* ）。人生活在两种基本关系中间：我-它（人与周围事物）关系和我-你（人与人）关系。在他看来，人与人之间的关系不过是人与神“相遇”的反映，人是可以与神对话的。《不列颠百科全书》称他对20世纪人类生活产生了深远影响。

[119] 疑病的英文很形象，worried well，没病的人总怀疑自己得了病（特别是可怕的大病，如艾滋病），那样的人仿佛生活在忧虑的深井里。这个词虽然不是严格的医学术语，但在专业文章里已经广泛流行了。

[120] 原初场景（primal scene）是婴儿被大人诱惑的场景，是弗洛伊德假定的一类精神病根源。

[121] 雅各布（Francois Jacob，1920—）是法国细胞遗传学家，1965年“因有关酶和病毒合成的遗传调节方面的发现”而与Andre Michel Lwoff和Jacques Lucien Monod共享诺贝尔生理学或医学奖。这句话是他在《苍蝇、老鼠和人》（*Of Flies, Mices, and Men* ）一书的最后说的。

[122] 坎德尔（Eric Kandel，1929—）出生在奥地利，现在是哥伦比亚大学神经生物和行为科学研究中心主任。他长期通过无脊椎动物海兔来研究记忆的分子生物学机制，获2000年诺贝尔生理学或医学奖。最近，他和他的伙伴们发现一种新的与学习性恐惧有关的脑电活动，有可能使我们以分子遗传学的方法来治疗那种恐惧。

[123] Woody Allen是美国有名的导演和演员，《安妮霍尔》（Annie Hall）1977年获Oscar最佳影片奖，这是一部自传性的影片，却没有完整的故事情节。主人公艾维辛格和安妮霍尔在网球场结识，然后相恋、同居，最终分手。艾维既缺乏安全感，又经常在事物两个极端之间徘徊，永远在自嘲和自省之间，在嘲笑别人时，自己也犯上同一类错误，讽刺了美国知识分子的犬儒态度。

[124] Lourdes是美国有名的医院，纽约有Lourdes医院，新泽西还有Lourdes医疗中心。

[125] 让病人暴露在各种刺激环境下，在反复暴露中取得适应，从而消除应激反应，是治疗恐怖症和强迫症等神经症最常用的行为疗法。暴露疗法适用于绝大多数焦虑性障碍，对抑郁症是否有效，还存在争论。

[126] 斯特拉奇（1880—1932）的《维多利亚名人传》（1918年出版）为英国维多利亚时代的四个名人（Cardinal Manning, Florence Nightingale, Dr.Arnold, General Gordon）做了新传，让名人回到凡人中间，也使传记回到文学艺术。这是一部影响20世纪传记文学的经典著作，作者自己写的前言成了这种“新传记”的宣言书。

[127] Valery Ryumin是俄罗斯宇航员。美国《发现》杂志2001年5月号在《我们能到达火星而不

发疯吗？》一文中也引用了这句话。

[128] 2000年，美国Charmed技术公司在伦敦互联网世界大会上让一个女模特展示了一种可穿戴的电脑：她头戴显示屏，腰间挂着主机，手上拿着无线鼠标。

[129] 这是一本在日本、韩国和中国都很畅销的书（中译本，盛海燕、刘雪芹、张进译，中国友谊出版社，2000）。书的最后，作者让我们记住女作家、“精神美女”艾略特（George Eliot）的话：“让我们赞美和尊敬神圣的形式美！让我们在男人、女人和孩子身上、在自己的花园和房子里培育美直到尽善尽美的程度。但同时也让我们热爱另外一种美：这种美的秘密不在于比例，它存在于人类深深的同情心。”

[130] 积木式的方法就是在不同的层次和对象上研究一个问题，然后把研究结果综合起来，希望它能最后解决那个问题。计算机技术的模块组合，就是一个典型的例子。

[131] 20世纪，很少有像托马斯（1913—1993）那样能沟通科学与文学的医生。他写过很多科学趣味浓厚的蒙田式的散文，有名的如*The Lives of a Cell—Notes of a Biology Watcher*，《第一推动》丛书已经有了《细胞生命的礼赞》。

[132] Infectious diseases过去通常译作“传染病”，现在认为传染病（communicable disease或contagious disease）只是感染性疾病的一类，应该放在感染性疾病的大环境下来研究。

[133] 中国在宋真宗（998—1022）时大概就知道种牛痘预防天花了，不过要留下疤痕。1796年5月14日，英国乡村医生詹纳（1749—1823）从一个挤奶姑娘的痘疮中取出痘浆，接种到一个8岁男孩的胳臂上——天花疫苗就是这样发现的。

[134] 西莫威斯（1818—1865）是匈牙利产科医生，证明了产褥热源于不清洁的医生的手和器械。通过消毒，产褥热的死亡率显著降低了。但是他的发现受到当时保守势力的嘲笑，最后因神经错乱而自杀了。李斯特（1827—1912）是英国外科医生，也是微生物学家。1865年，他从一个化学家朋友处得到一些石碳酸，用来浸泡绷带，这是外科手术消毒杀菌的第一次试验。现在常用的“李斯特灵”消毒水，就是以他的名字命名的。埃尔利希（1854—1915）是化学疗法的创始人，1908年，他与俄罗斯Ilya Ilyich Mechnikov一起，因为在免疫研究上的贡献获诺贝尔医学或生理学奖。1910年，他发明了治疗梅毒的洒尔福散（Salvarsan）。弗莱明（1881—1955）是英国伦敦圣玛丽医院的细菌学家，他在1928年发现了青霉素（历史上第一种抗生素），但他没有方法把它提纯，到了1943年青霉素才开始用于临床。

[135] 这句话的意思是说，在同一个环境下存在的相互矛盾的两个因子，就像博弈游戏中的双方；协同进化的过程就像博弈游戏的过程。

[136] 卡波西肉瘤是发生在皮肤上的紫色斑点，30%到50%携带疱疹病毒8的HIV（爱滋病病毒）感染者最终将患卡波西肉瘤。这种病毒多见于同性恋者。

[137] 半个多世纪以后（1966年），劳斯获得了诺贝尔生理学或医学奖，因为他“发现了第一个

可引起动物实质肿瘤的病毒，从而开辟了一个研究肿瘤与病毒关系的新领域——这领域不但对了解肿瘤成因，而且对了解正常细胞如何转变成肿瘤细胞也十分重要”。

[138] HIV来自黑猩猩的SIV；1992年人们在巴尔干山探险时感染马堡病毒；1993年美国流行汉塔病；1995年扎伊尔爆发埃波拉出血热；1996年英国爆发疯牛病；1999年在美国发现西尼罗河病毒（1937年在乌干达尼罗河西部地区发现的一种奇怪病毒，主要由鸟类携带，通过蚊子传播。1999年，美国有人因为这种病毒死亡，才引起大家的关注。2002年蔓延40个州，2003年8月又在美国扩散）；2003年中国流行SARS……最近，科学家们发现，我们的免疫系统的第一道防线有可能将Ebola病毒引入我们的细胞（据2003年6月26日Nature消息）。所有这些例子似乎都是在人与自然的新关系中产生的，关于它们的未来，我们现在的任何预言可能都没有意义。

第一推动丛书：综合系列

真理与美

[美] S．钱德拉塞卡 著

杨建邺 王晓明 译

The Polytechnique Series

Truth and Beauty

S. Chandrasekhar

THE FIRST MOVER

湖南科学技术出版社

版权信息

真理与美

著者：【美】S.钱德拉塞卡

译者：杨建邺　王晓明

责任编辑：李永平　吴炜　戴涛　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794413

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
前言
第1章 科学家（1946）
第2章 科学的追求及其动机（1985）
第3章　诺拉和爱德华·赖森讲座　莎士比亚、牛顿和贝多芬：不同的创造模式（1975）
第4章 美与科学对美的探求（1979）
第5章　米尔恩讲座　爱德华·阿瑟·米尔恩和他在现代天体物理学发展中的地位（1979）
第6章 纪念A.S.爱丁顿诞辰一百周年讲座（1982）
第7章 K.史瓦西讲座广义相对论的美学基础（1986）
附录
附录 寻求秩序——钱德拉塞卡对黑洞、蓝天和科学创造力的思考
译后记

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序　一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

前言

S.钱德拉塞卡

1986年12月8日

本书收集的是我的7篇演讲，它们反映了我对于科学研究的动机和科学创造模式的一般观点。第一篇演讲是40年前做的（具体情况我在下面还要讲到），其余6篇是在1975年之后的10年中做的。正因为前后相隔几十年，所以它们显示了一个科学家态度的变化。（或成熟？）

这些演讲都做过精心的准备，在内容的细节以及措词上也做过认真的考虑。事实上，它们都是在一些重要的讲座上宣读的；收集在本书中时原稿未做改动，只删掉了一些开场白。

Ⅰ

这些演讲大致上有两方面的内容。前4篇主要阐述美学和动机的问题。其余冠有米尔恩、爱丁顿和史瓦西讲座的3篇演讲，虽然其部分内容是介绍他们各自的经历，但也都间接地谈到了上述一般问题。特别是在卡尔·史瓦西讲座的演讲中，主要讨论的是广义相对论的美学基础，它是前面《美与科学对美的探求》讨论的继续。

Ⅱ

从1946年做《科学家》的演讲到1976年做《莎士比亚、牛顿和贝多芬：不同的创造模式》的演讲，中间相隔30年。前面我已说过，这是由特殊环境形成的。科学家一般都认为，科学追求的动机或这种追求的美学基础，是不值得认真讨论的；而且对认真讨论这些问题的科学家，他们也往往持怀疑态度，甚至不屑一顾。我在1945年大致上也持有这种观点。但是，当时任芝加哥大学校长的哈钦斯（R.A.Hutchins）却给我写了一封信，邀请我在他组织的系列讲座中做有关《科学家》的演讲，他

在信中解释道：

> 这次系列讲座的目的是激发大学生的批评能力，使他们了解什么是优秀的工作，引导他们尽力把各自的工作做好。希望每位演讲人谈谈他自己从事本行工作的体验，通过阐述其特性、总结其目的以及解释其技巧，来说明各自工作的价值。

开始，我不大愿意接受邀请，因为对这些问题我没有认真思考过。此外，哈钦斯邀请的其他演讲人的名单中还有赖特（F.L.Wright）、勋柏格（A.Schoenberg）、恰卡尔（M.Chagall）、冯·诺伊曼（John von Neumann）[1]这些赫赫有名的大人物，使我心虚、怯场。想想看，谁看了这张名单不会吓一跳。但那时我还很年轻，无法抗拒一位大学校长的权威，我只好硬着头皮去思考那些我当时还很生疏的问题。

当我再次看40年前我的讲稿时，我感到有些话我今天不会说，或者说法有些不同。但我还是把它收进了这本书，因为把1946年的演讲与1985年的演讲《美与科学对美的探求》放在一起，也许有助于读者更好地判断一位科学家对问题的观点如何随时间变化。

Ⅲ

从时间顺序上看，《科学家》演讲之后是1975年《莎士比亚、牛顿和贝多芬：不同的创造模式》的演讲。1974年，由于生病我不得不疗养了半年，这使我有了一次难得的机会，可以专心致志地思考一些我从未认真思考过的问题。半年的学习、思考和研究，不仅为我即将做的演讲提供了基础，而且使我对美的敏感性在科学素养中起的作用，产生了持续的兴趣。对广义相对论的数学方面研究得越深入，就越是加强了我的这一兴趣。（我应该补充一点，我所发现的新事实或新见解，在我看来并非我的“发现”，而是早就在那儿，我只不过偶然把它们拾起来罢了。这看来有点奇怪，但这是真的。）

Ⅳ

1975年以来的所有演讲中，用来阐述我的观点的一些相同的“故事”，在不同的背景和不同的地方出现，但有两条相互交错的线索把它们串联起来。一条线索是关于在科学中对美的追求，另一条线索是关于

艺术和科学中不同创造模式的起源，这是我在1975年讲演中明确提出来的。这两种创造模式的明显差异，在我们讨论一位艺术家的工作和一位科学家的工作时，可以清楚地看出来。在评论一位艺术家时，我们常常把他们的工作区分为早期、中期和晚期；这种区分一般表示出作家成熟程度和认识深度的不同。但在评论一位科学家时，却往往不能这样。对科学家往往是根据他在思想领域或实践领域做出的一个或几个发现的重要意义来做出评价。一位科学家最“重要”的发现往往是他的第一个发现；相反，一位艺术家最深刻的创造多半是他最后做出的。这种明显的差异至今仍然令我感到迷惑不解。

最近，我突然悟出了一点道理，也许有助于认识这种明显的差异，我不妨简略地说一下。16世纪和17世纪科学家的目的与现代科学家有明显的不同。牛顿是最突出的例子。在大瘟疫时期他避居于家乡伍尔兹索普，这一期间他发现了万有引力定律和其他一些定律。大约20年之后，在哈雷的请求下他才重新写出开普勒第一定律的推导，但他没有就此打住，他甚至也不满意他随后做的演讲《论物体的运动》。不写完全部《原理》他是不会罢手的：他写这本书的速度和连贯性，在人类思想史上真是无与伦比。从现有的认识水平来看，牛顿的拼搏在一个方面给人们以启迪，那就是他并不急于宣布他的发现；他想完成的研究远不止这一个发现，他似乎要把他的发现放在整个科学领域之中，而且他认为科学是一个整体，是一个他有能力建成的整体。在牛顿所处的时代，这种科学观比较普遍，例如开普勒在给出行星运动定律后，他本可心满意足，但他却决定写一本《新天文学》。伽利略也是如此，他在做出他的一些伟大发现后并没有停步，他显然认为他必须写出《关于两种新科学的对话》。后来，拉普拉斯和拉格朗日继承了开普勒、伽利略和牛顿的这一传统。

当然啦，如果现在一个正常的人还去刻意模仿牛顿、伽利略和开普勒，别人一定会取笑他，认为他闲着没事干。但是，这些范例表明，以巨大的视野作为科学的目的在科学史上确实存在过，而现在科学的目的则没有往日那么宏大。现在的科学目的逐渐转向强调改变科学方向的发现上，这种改变也许是大势所趋，不可避免。与伏打、安培、奥斯特和法拉第名字相关联的一些发现，必然先于麦克斯韦的综合；它们各自需要不同类型的努力。无论如何，强调“发现”的倾向仍在继续，而对在科学发现中如何理解取得科学成就的主要因素，则进一步突出和强化了这种倾向。用一个简单的框架把某人的想象综合起来，即使在有限的范围

里，也已经失去了价值。例如，我们不会向爱因斯坦提出这样一个问题：在发现他的引力定律20年后，他有没有设想（或感到能够）写一本像《原理》那样的书来阐述广义相对论。

假如16世纪和17世纪的伟大科学家对科学追求的目的在今日仍然通用，那么艺术家和科学家在创造模式上的差别，也许就不会出现了。这种看法正确吗？

我还想补充一点，在确定哪些演讲该收进这本集子中时，我与妻子拉莉莎（Lalitha）进行了深入的讨论。她的鉴别能力和毫不逊色的洞见，对本书最后定稿起了重要作用。我还应该感谢她对我不断的鼓励。

第1章　科学家（1946）

首先我得承认，让我作为“思维的作用”这个系列讲座的一个撰讲人，我感到担忧。因为讨论科学家的创造力这样的问题，必然涉及广阔而又全面的知识，而我深深感到我在这方面不是行家，可能讲不好。尽管我对把我作为这个系列讲座中科学家的代表是否合适感到疑虑，但对选择天文学和天体物理学作为精密科学的代表，我却没有丝毫疑虑。因为在所有精密科学的学科中，天文学最具综合性。它需要综合各个不同时期的学术成就，以便在实践中逐步完善。另外，在所有科学中天文学占有独特地位，诺伊格鲍尔（O.Neugebauer）[2] 曾经说过：

> 自从罗马帝国衰亡以来，天文学是所有古代科学学科中唯一完整流传下来的分支。当然，在罗马帝国残存的地域内天文学研究的水平下降了，但天文学理论与实践的传统却从来没有丢失。相反，印度和阿拉伯的天文学者改进了希腊三角学的笨拙方法，新的观察结果不断地与托勒密的观察结果加以比较，等等。人们只有将这种情形与希腊数学的较高分支的完全失落这一情形加以对比，才能认识到天文学是联系现代学科与古代学科的最直接环节。的确，只有不断地参考古代的方法和概念后，人们才能理解哥白尼、第谷·布拉赫和开普勒的著作，但是，我们要想理解希腊人有关无理数的理论和阿基米德的集合方法，那只有现代科学家在新发现它们后才可能。

这个系列讲座的发起人要求每个演讲者通过阐述其特性、总结其目的以及解释其技巧，来说明他所从事的艺术或职业的价值。在我开始讨论这些问题之前，我想提请大家注意并牢牢记住自然科学的总体分类，即自然科学分为基础科学和导出科学（derived science）两类。请大家注意，我没有在“理论科学”和“应用科学”之间做出什么区分。对于后者我不打算讨论，因为我不相信在刻意追求科学的应用中，会发现科学的真正价值。因此我将只讨论通常所说的“理论科学”，我想要大家注意的是，我的分类正是将理论科学分为基础科学和导出科学两部分。尽管无法对基础科学和导出科学给出准确或鲜明的定义，但这种分类确实存在着，并且通过我要枚举的例证，它将表现得越来越清楚。广义地说，我

们可以认为基础科学试图分析物质的终极构成和基本的时空观；而导出科学所关心的是，利用这些基本概念将自然现象的各个侧面条理化。通过这样的叙述，有两点是很清楚的：第一，这种分类依赖于在某一特定时间内科学所处的状态；第二，在分析自然现象时，可能确实存在着不同的层次。例如，大量的现象能够从牛顿定律有效的领域中找到直接和自然的解释。然而，其他类型的一些问题就只能从量子理论中获得答案。既然存在如此不同的分析层次，那么肯定存在一些判据，利用这些判据，我们就能确定，在什么情况下哪些定律是适用的，哪些是不适用的。

至于讲到分类本身，我认为最好的例子莫过于卢瑟福（E.Rutherford）发现α粒子的大角散射。他做的这个实验非常简单。用某种放射性物质发射出来的高能α粒子轰击一层薄箔时，卢瑟福发现α粒子有时被完全弹了回来——这种完全弹回的粒子很少，但确确实实存在。在他晚年（1936年）回想这种现象时，他说：“这是我一生中所遇到的最难以令人置信的事。”他还这样描述过他当时的反应：“其难以置信的程度就像用一发15英寸的炮弹射击一张卫生纸，炮弹反弹回来并击中炮手。”他还写道：

> 经过仔细思考，我马上意识到这种反方向的散射肯定是出自某种单一的碰撞。经过计算我发现，除非重建一个原子模型。在新模型中原子的绝大部分质量都集中在某个很小的核上，否则不可能得到这种数量级的散射结果。正是从那时起，我认为原子有个很小但很重的带电质心。我发现，某一给定角度的散射粒子数与箔厚成正比，与核电量的平方成正比，与粒子速度的四次方成反比。这些推论后来被盖革（Geiger）和马斯顿（Marsden）用一系列漂亮的实验证实。

作为所有学科基础的原子核模型就这样产生了。一个唯一的观测和对此所做的正确解释，竟导致了科学思想的革命，这在科学史上也是无与伦比的。

我认为，查德威克（J.Chadwick）发现中子一事也属同一种情形。人们现在相信，中子和质子是所有原子核的基本组成成分。

但是我们不能只根据这两个例子就认为，所有有关基础科学的事例

只能在原子物理学的领域中才能找到。事实上，能被称为“基础”定律的首例起源于天文学，我指的是开普勒（J.Kepler）的发现。开普勒对第谷·布拉赫（Tycho Brahe）的大量观察结果做了长时间和耐心的分析后，终于发现了行星运动的定律。后来，开普勒定律又导致了牛顿（I.Newton）著名的万有引力定律，而牛顿万有引力定律两百多年来一直在科学舞台上起主导作用。过一会我还会回过头从不同的角度来讨论这个问题，但这个例子足以说明，只有在万有引力的领域里天文学才能直接引出具有基础性的结论。还有一个例子可以说明这件事实，水星的实际运动轨道与根据牛顿定律预测的轨道之间存在着细微的偏差，该偏差指出了且随后证实了由广义相对论蕴含的对时空观的根本变革。这一事实进一步说明了上述“天文起源”（即“基础定律首例起源于天文学”）问题。哈勃（E.P.Hubble）发现银河系外星云正在远离我们而去，其远离的速度与它们跟银河系的距离成正比，同样，这一发现颇有可能导致我们基础概念的进一步修改。

我上面所举的几个例子，或许表明了科学的真正价值存在于能直接导致我称之为“基础”进展的追求之中。事实上，有许多物理学家真的接受了这种看法。例如，一位很杰出的物理学家曾经对我说，我早就应该是一个真正的物理学家了。显然，他对于我特别偏爱天文学的事情感到担忧，同时也想鼓励我。我认为，这种态度代表着一种对于科学的真正价值的误解，并且，科学史也会对这种态度提出异议。从牛顿时代至20世纪初，整个动力学和由它演绎出的天体力学都完全是在对牛顿定律的结论做扩充、完善和计算。哈雷、拉普拉斯、拉格朗日、哈密顿、雅可比、庞加莱——他们都乐于将他们科学生涯的大部分精力用在这件工作上，也就是说，用于推广一门导出科学上。对于导出科学的嘲笑，意味着否定了这些人如此严肃认真追求的价值观。这在我看来，简直是荒谬透顶得不值一提。公正地说，基础科学和导出科学之间很明确地存在一种互补关系。基本概念的有效程度，与它们能分析的自然现象的范围大小成正比。如果限制这些概念的有效范围，我们就会发现其他定律的应用将比我们用过的定律更加普遍。从这种观点来看，科学永远是一个形成过程，正是在这种共同努力去分享科学进展的过程中，科学的价值才能得到实现。我想有了以上一些看法，我就能以一种更正式的方式，叙述我所认为的科学的真正价值，这种价值也正是一个科学家在他的实际工作中所追寻的。

科学的价值在于对自然的一致性的不断完善的认识之中。事实上，

这仅仅只是意味着这些价值的获得，或大或小地扩大或者等量地限制了人们关于物质及时空概念的适用范围。换言之，科学家期望在他们的追求中，能不断地扩大某个基本概念的适用范围。在这样做时，科学家试图发现这些同一概念是否存在着某些限制，并试图形成范围更宽和适用性更大的概念。科学家所追求的这些价值，包含在我将讨论的三种不同形式之中，这三种不同形式的标题是："基本定律的普适性"、"根据基本定律所做的预测"和"由基本定律做出的证明"。

我将通过实例分别阐述它们。

基本定律的普适性

通过讲述万有引力定律的普适性，在某种程度上能很好地描绘出引力定律是如何获得普适性的。

人们早已发现，地球上所有物体均受到一个指向地心的引力作用。然而这种引力能够影响到多大范围呢？它能影响到月亮那么远的地方吗？牛顿向自己提出了这些问题，并且他回答了它们。伽利略已经证明，匀速直线运动和静止都是物体的自然状态，偏离这种自然状态需要力的作用。假定月亮不受任何力的作用，它将脱离轨道而沿轨道的瞬时切线方向离去。如果月亮的运动是由于地球引力形成的，那么这种引力的作用就是把月亮从瞬时切线方向拉到轨道上运动。由于月亮绕地球转动的周期和距离都是已知的，所以很容易算出月亮在1秒钟内由切线落下的距离。将这个值与自由落体的速度比较之后，牛顿发现两者之比为1:3600。又因为月亮到地心的距离是地球表面上物体到地心距离的60倍，这就意味着存在一个与距离平方成反比的力。

牛顿向自己提出的第二个问题是：引力的这种性质到底在多大程度上有效。特别是太阳是否也有类似的力使行星做轨道运动，就像地球引力使月亮做轨道运动一样？这些问题的答案可在开普勒定律找到。牛顿指出：开普勒第二定律——行星在相同时间内掠过相同面积——意味着存在一有心力，即指向太阳的一种力；开普勒第一定律——行星轨道为椭圆且以太阳为椭圆的一个焦点——是引力平方反比定律的一个结论；最后，若同一定律对各个行星均成立的话，那么，行星运动的周期和距离的关系就在开普勒第三定律中得到表述。牛顿就是以这样的方式阐明他的万有引力定律，即宇宙中任一粒子对其他任何粒子都有引力作用，

其大小与它们之间距离的平方成反比，与两粒子的质量成正比。应该注意到在这个公式的描述中用了“宇宙”这个词，这就很清楚地表明了该公式的重要性在于它的普适性。

再举一个与观测有关的例子。威廉·赫歇耳（W.Herschel）根据他对距离很近的恒星对的研究，于1803年宣布：在某些时候恒星对是双星，它们相互绕着对方旋转。赫歇耳还进一步阐明，它们的表观轨道是椭圆，而且开普勒面积定律同样可以适用。换言之，这种观察结果使得万有引力定律的适用范围从太阳系扩展到了遥远的恒星。我们现在很难想象得出赫歇耳的发现，对他同时代人有多么大的影响。

自从牛顿定律公之于众后，天文学中众多的进展都与牛顿定律在太阳系运动中的应用有关。牛顿本人就得出了很多的重要结论。这里只讲两个例子：其一，他正确地解释了海洋的潮汐现象；其二，他还正确地解释了距他两千多年之前喜帕恰斯（Hipparchus）就发现了的岁差现象。

将牛顿定律运用到整个太阳系是一项极其艰巨的任务，它耗尽了许多科学巨匠毕生的精力，如拉格朗日、拉普拉斯、欧拉、亚当斯、德劳雷（Delaunay）、希尔（Hill）、纽康姆（Newcomb）以及庞加莱等。

我前面已经讲过，用牛顿定律不能完全解释水星的运动。水星实际运动轨道与牛顿定律计算出来的轨道有小小的偏离，这种偏离体现为一种整个轨道的缓慢进动，该进动速率比用牛顿定律计算出来的速率要超出一点点，即100年只有42弧秒。现在，用爱因斯坦的广义相对论似乎已经能圆满地解释水星的这种进动。

现在，牛顿定律仍能有效地运用到天文学众多的领域里。其中最新的领域是将整个银河系的运动作为一个整体进行研究，这个动力学的新分支称为“天体动力学”。其发展极为迅速，有着广阔的前景。下面有几个地方我还会涉及这一领域。

让我暂时撇开自然界定律普适性的经典例子，来看一个更新颖，在某种程度上更令人惊讶的例子。核衰变现象（通常叫作“原子裂变”）在近年来已得到广泛的研究，对于1946年的听众来说，这件事毋庸赘言。利用研究结果，贝特（H.A.Bethe）在几年前宣布：涉及碳和氮的某些核衰变可间接合成由四个质子组成的氦核。他还进一步指出，根据前不

久天体物理学家推导出来的太阳内部的情况，再利用在实验室发现的反应截面，我们现在能相当满意地解释太阳能的来源——这又是一个许多不同类型的研究综合起来解释某一现象的辉煌例证。

我们再看另一个例子。1926年，当费米（E.Fermi）和狄拉克（P.A.M.Dirac）将统计力学定律应用于电子气（electron gas）时，他们不得不对这些定律做些修改，并且证明在高密度和（或）低温度的情形下，经典定律将出现偏差。这种偏差的性质表现在：根据经典定律，压力正比于密度和温度。若在给定的温度下增加密度，偏差就会逐步表现出来，即随着密度增加压力迅速增加，并最终变成了只是密度的函数。这种状态称为电子气的简并态。这些新的定律在金属学理论中有着广泛的应用，并且有极大的实用价值。然而这些新定律的最初应用却发生在天体物理学中。R.H.福勒（Fowler）利用费米—狄拉克气体定律阐明了类似天狼星的伴星这种高密度恒星的结构。通常称之为白矮星（white dwarfs）的这类高密度恒星，其密度数量级达每立方英寸几吨。最特别的例子是几年前由G.P.克尤帕尔（Kuiper）所发现的一颗恒星，据估计其密度达每立方英寸620吨。福勒立刻认识到，在此情形下，根据费米—狄拉克统计学，电子一定处于简并态。有了福勒的发现，人们才可能算出白矮星的构成成分。

我个人对于白矮星的结构这个题目比较感兴趣，如果对这个题目讲得多了一点，请大家谅解。随着福勒论点的推广，人们很快发现费米—狄拉克定律需要做进一步修改以便解释这样一个事实，即在高密度的白矮星中必然有相当数量的电子以接近光速的速度运动。当考虑到如此高的速度并做了修正之后，人们就发现高密度星体的质量存在着一个上限。该上限大约是1.4个太阳质量。该上限出现的原因是若超过该质量就不会存在稳定的平衡组态。认识到这种质量上限又引起了很多关于星体演变的有趣问题。超新星现象的出现与此有某种关系是完全有可能的。这方面的研究我不能再进一步讨论下去了。我之所以提到这些问题，就是想要大家注意到：某些基本定律的有效范围是不断扩大的。

我上面所给出的三个例子，都是讨论同类定律的适用程度。但是，有时我们将同一类思想应用到各种问题中去，而这些问题乍看起来可能毫不相关。例如，用于解释溶液中微观胶体粒子运动的基本概念同样可用于解释星群的运动，认识到这一事实是令人惊奇的。这两种问题的基本一致性——它具有深远的意义——是我一生中所遇到的最令人惊讶的

现象之一，对此我想多讲几句。

“布朗运动”现象是英国植物学家布朗（R.Brown）在1827年发现的。当他观察悬浮在水中的微粒（他用的是花粉）时，他发现这些微粒永远不会静止下来，处于一种不停地骚动的状态。现在想来似乎可笑，起初这种不停的运动竟被认为是花粉的生命活动引起的，但是布朗马上就指出这种解释是不可能的。因为即使是从埃及斯芬克斯石像上取得的细微尘埃，也具有同样的行为。现在我们知道，布朗运动起源于胶体微粒与它们周围的液体分子的碰撞。既然最细小的胶体微粒也要比单个的分子重几百万倍，显然单一的碰撞几乎不会对胶体微粒产生任何影响。但大量碰撞的总体效应是可观的。令人惊诧的是用于研究布朗运动的同样方法，也适用于研究像昴星团这样的星群运动。我们可以这样做的原因是：当星群中的两颗星擦身而过时，每颗星体运动的方向和量值都发生了变化。由于星体间的作用力与距离的平方成反比，作为单个效应来说星体的运动受到的影响很小，但同样地由于大量的这类交遇，其累积效应就产生了可观的变化。很明显，这与布朗运动是类似的，所以星群运动理论能够随着布朗运动理论的发展而发展。而且，星群运动理论比胶体微粒运动理论更完备地描述了布朗运动的特征。我还要指出的是，正是由于这种理论的发展，我们才能在总体上预言星群的演变和宇宙的时间尺度。

根据基本定律所做的预测

现在我要谈到科研的一个侧面，即根据其他证据推出的定律来做预测，以及对这些预测所做的证实。

我认为，在近代所做的且随后被证实的预测中，最令人瞩目的要算是哈雷的预言了。1705年，爱德蒙德·哈雷（E.Halley）向皇家学会做了《彗星摘要》（Astronomiae Cometicae Synopsis）的专题报告。在这份经典的论文中，哈雷仔细研究了从最早年代甚至牛顿时代有关彗星的各种记载。接着，根据牛顿原理，哈雷对从1337年到1698年间做过专门观察的24颗彗星进行了抛物线性的计算。这份论文的准确性和完备性可以说是达到了无可挑剔的程度，对人类知识做出了既有纪念意义又令人回味无穷的贡献，读起来叫人爱不释手。正是在这篇论文中，哈雷想到了这种可能性，或者说或然率，即彗星的运动轨迹可能是极扁的椭圆而不是抛物线。在后一种情况下彗星来自无穷远处，也将归宿于无穷远处。

然而，在前一种情况下，彗星就是太阳系的成员了，经过漫长的若干年，它们将重新出现。正因为有这种可能性，哈雷才做了大量的计算工作。这样，如果出现一颗新的彗星，可将它的轨迹与已计算出的轨迹相比较，我们就可能确定它是否是曾出现过的彗星。哈雷还说，许多迹象使他确信，1531年的那颗彗星与1607年观察到的彗星以及1682年他本人亲自观察过的彗星是同一颗，他还认为大约在1456年看到的那颗彗星也就是这同一颗彗星。随后他写道："由此我很有信心地大胆预言，这颗彗星将于1758年重新出现。"这就是彗星中最著名的哈雷彗星的起源。哈雷没能看到这颗彗星再次出现就去世了，但它确实在哈雷所预言的那一年出现了，并且在此之后又出现过两次。

预言随后被证实的另一个新近的事例是狄拉克关于正电子的预言。1928年，狄拉克灵感突发，写出了一个有关电子的方程。这个方程预言的许多事情都与实验吻合，但该方程还预言电子应该有负能态——这可真是前所未闻！然而，狄拉克与以往一样坚信他的方程是正确的，他断定存在着负能态。为了克服所有电子都坠入负能态并在人们周围产生一个奇妙世界这样的难题，狄拉克提出了他的设想：在通常情形下，所有的负能态都被填满了，极少数带有正能的剩余电子不能进入负能态，通常情况下事实确实如此。尽管如此，在某些条件下负能态的电子能够被激发到正能态，这样就产生一个电子并在无限分布的负能态中产生一个空缺，正是这个无限分布的负能态中的"空穴"会表现得如同一个完全可察觉的正能粒子一样，不过带正电荷而已。这个空穴就是正电子，狄拉克假设的现象就是电子对的产生。狄拉克甚至还建立了一套有关这种电子对生成的概率理论。大约三年后，所有这些预言都得到了证实，这使他更加坚信他的方程是绝对正确的。

预言被证实这一类情况中，我想讲的第三个也是最后一个例子是爱因斯坦关于引力场中光线会发生弯曲的预言以及对该预言的证实。在讲述这个故事时，我要摘录证实预言的主要人物爱丁顿（A.S.Eddington）在一次演讲中的几段话：

> 在我的天文学生涯中，我能想起的最令人激动的事件要算在1919年的日食观察中，证实了爱因斯坦有关光线发生弯曲的预言。当时的情况是很不寻常的。虽然于战争期间的1918年开始制订了计划，但直到出发前11个小时我们还在怀疑这次考察能否成行。但1919年的日食太重要了，不能错过这千载难逢的良机，因为这次具

有极好的星场——任何之后的考察都不会有这样好的时机。已故皇家天文学家弗兰克·戴逊（F.Dyson）爵士在格林尼治组织了两支考察队，一支赴巴西的索布拉尔，另一支赴西非的普林西比岛（普林西比岛考察队由爱丁顿负责）。显然，要想在停战之前让仪器制造商制造一些观察用的仪器是不可能的。由于考察队得于2月份出发，所以准备工作极为仓促。巴西队在日食那天天气异常好，可惜碰到了一些情况，所以他们的观察结果几个月后才处理出来，但最后他们还是提供了关键性的证据。我当时在普林西比岛。日食那天下起雨来，满天乌云，大家都几乎完全失望了。接近全食时，太阳隐隐地显露出来。我们抱着一线希望执行了原订计划。一定是日全食结束之前乌云变薄了一点，因为尽管有许多底片报废了，可我们仍得到了显示出要找的星象的两张照片。将它们与太阳在别处时同一星场所摄的照片加以比较，有明显位移，这表明星光在掠过太阳时，光线的确发生了弯曲。

这个问题有三种可能性：其一，可能根本就没有什么弯曲现象，即光线不受引力场的影响；其二，可能存在一种“半弯曲现象”，即光线受引力场影响服从牛顿定律而发生弯曲；另一可能是服从爱因斯坦而不是牛顿定律的全弯曲现象。我记得戴逊向已知道这些主要思想的柯丁罕（Cottingham）解释这一切时说，光线弯曲得越厉害其结果就越令人激动。柯丁罕问：“如果我们得到双倍的弯曲会怎么样呢？”戴逊说：“那样的话爱丁顿就会发疯，你就只好一个人回家了。”

当时就对照片做了测算，这不仅仅是急不可耐，而是怕在回家途中发生什么不幸。于是对成功的两张照片中的一张立即进行了检试。结果得到的数值从天文学标准来说已经完全足够了，所以一张照片实际就可以确证一切了，尽管还会从其他方面寻求进一步的证实。日食后第三天，当算完最后一个数据时，我意识到爱因斯坦理论经受住了实践检验，新的科学思想观必将受到广泛承认。柯丁罕也将不会是独自一人回家的。

由基本定律做出的证明

我现在讲讲科研的第三个特性，在某种意义上，这一特性介于我已经讲述的两个特性之间。

18世纪，唯心主义哲学家贝克莱大主教和他的追随者宣称：太阳、月亮和星体只不过是“我们头脑中的感知”，探索诸如星体构成之类的问题毫无意义。但只过了几十年，即在1860年基尔霍夫（G.R.Kirchhoff）宣布了具有重大意义的有关夫琅和费线的化学解释。他指出，夫琅和费线表明在太阳的大气层中，人们所熟知的一些金属元素以炽热的蒸气形式存在。从那时起，谈论星体的构成再也不是痴人说梦的事，而是具有重大实际意义的问题。

在此后的80多年中，人们将实验室和天文观察得到的无数光谱，做出了几近完备的解释，其时间之短和任务之艰巨，真令人难以置信。有关研究这些问题的故事组成了科学历史和科学方法中最富于浪漫色彩的章节之一。当然该章节中的大部分内容都不能脱离50多年来物理、化学和天文学的发展史。如果我从这个大领域中挑出两个细节来专门讨论，那并不是我过于强调它们的重要性或意义，而只是我碰巧对它们特别感兴趣而已。我想谈及的是氢之后的两个最简单的原子，即带两个电子的原子：氦和负氢离子。

首先谈谈氦。1895年3月之前，人们只知道氦是太阳色球层的一种色球元素。1868年8月，法国天文学家詹森（Jansen）在日全食中探测到了氦的存在。詹森观察到的现象是：全食时，当太阳在喷射炽热气体的瞬间所获得的色谱中，在众所周知的钠线附近存在着一条波长为5876的明亮黄线。起初人们认为该线可能是钠引起的。但洛克耶爵士（Sir Norman Lockyer）首先意识到这种解释是不正确的，而且这条新线与当时已知的地球上的任何元素谱线都不吻合。因此，他断定出现了一种新元素。又因为该元素是在太阳中探测到的，所以他称之为氦。1895年即四分之一世纪之后，著名的化学家威廉·拉姆赛爵士（Sir William Ramsay）在研究某些铀矿物产生的气体时，检查了这些气体的谱图。他发现在谱图中有一明亮黄线正处于上述太阳谱图的氦线位置上。进一步的研究，确认了在两种情况下的谱线都是同一元素产生的。这样，首先在太阳上探测到的元素随后在地球上被分离出来了。

负氢离子的故事在某些方面同样引人入胜。由一个质子和两个电子组成的原子可以以自由态存在，这是贝特和海勒拉斯（Hylleraas）在理论基础上确认的事。贝特和海勒拉斯依据量子理论计算结果是如此明确肯定，以至无论是它的稳定性，还是这种原子在适当条件下以自由态存在的能力，都是不容置疑的。但迄今为止，在实验室中仍未分离出负氢

离子。不过威尔特（Wildt）前几年指出，负氢离子肯定会以自由态存在，而且大量存在于太阳大气层中。这就出现了一个问题："我们能否探测到它呢？"为了能探测到负氢离子，首先我们必须知道负氢离子吸收光线的方式以及这种吸收作用在太阳谱图中的表现。确定负氢离子如何吸收光线的理论问题竟然异常棘手，但是根本的物理学问题现在已经解决，人们可以相当肯定地预言，在太阳谱图中可能会观测到这种效应。这些效应的性质是如此清晰明了，并被观测如此充分地证实，所以可以毫不夸张地说：量子理论预言肯定存在由一个质子和两个电子组成的这种稳定的原子，很快会得到证实。

到此为止，我还只讲了科学家在他们的各自专业领域的实践活动中所做的探索和追求。在本文快结束时我想要谈一下科学家的动机。关于这个问题有几种不同的看法。有人认为科学家的动机源于他们有意识地或下意识地相信他们所做的一切，最终会给人们日常生活带来舒适，我不同意这种看法。有人坚称科学家必须总是有意识地将他们的工作与时代和社会的需要相结合，我也不赞成这种推论。有人认为科学家努力工作是因为他们对追求真理有一种"神圣的激情"或对于解开自然界的"奥秘"有一种"炽烈的好奇心"，这种看法我也不能接受。我不相信每天沉浸于工作的科学家，与放弃帝王生活而沉思对人生有意义的伦理和道德价值观的释迦牟尼之间会有什么共同之处。而且，我认为科学家与马可·波罗也不会有什么共同之处。

实际上科学家努力工作的具体和现实的原因是他们的那种愿望，即他们想尽自己最大的能力积极参与科学的进展过程。如果一定要我用一个字眼来描述激励科学家工作的主要动机，我就用"系统化"（Systematization）这个字眼。这听起来似乎太平淡无奇，但我认为它揭示了实质性的东西。从根本上说，科学家试图做的工作就是选择某一领域，某一方面或某一细节，来检验它们在具有一定形式和连贯性的总体框架中是否占有适当的位置；如果它们的位置不当，科学家的工作就是做进一步的探索以使它们占有适当位置。这种说法也许有点晦涩难懂，尤其是使用了"适当""总体框架""形式"和"连贯性"等字眼。我承认要定义这些字眼就如同要到艺术中定义美一样，但对那些熟悉自己研究课题的人来说，认识和欣赏这些术语并不困难。我不妨试着用两个简短的例子说明我的意思。

1896年，亨利·贝克勒尔（Henry Becquerel）发现了放射性现象。

关于放射性我们现在已经知道的内容是：有三种放射性系；当发生放射性衰变时会发射出一到三种不同的射线；放射性位移具有某种规律；存在着同位素和等量异位素；原子的自发衰变涉及新奇的理论；等等。可以想象，对于那些对上述事情一窍不通的人来说，放射性现象是多么复杂和变化多端。然而，1904年卢瑟福的《放射性》（Radioactivity）一书的第一版问世时，放射性现象的实质性问题就被揭示出来了。这个问题的解决，很大程度上归功于系统化地研究了能量、秩序和完备性——这些都是卢瑟福的特点。

再举一个例子：在第一次世界大战和20世纪20年代中，物理学家承担了揭开复合原子光谱之谜的极为浩繁的任务，如果不是有意识地做到了我所说的“系统化”，就不可能完成这项任务。20年代末确定的量子理论原理也是依这种方式才建立的。这个系列讲座的主办人曾明确表示过，希望每个演讲人讲讲自己的经历，因此我也不妨谈谈我所采取的工作方法。我的工作方法一直是：首先了解一个课题的已知情况，然后检查这些情况是否符合一般人们会关心的严谨性、逻辑条理和完备性的标准；如果不符合这些标准，就着手使之符合。在已有的学术成就上系统化，一直就是我的动机。我敢大胆地说，这的确是非常普遍的情况。无论如何，在我看来只有这样才能正常地进行科学研究，才能获得真正的科学价值。

我恐怕没有多少时间来讨论科学工作的另一个极重要的方面，即科学的集体合作性。在此，我只摘录卢瑟福的一段话：

> 任何个人要想突然做出惊人的发现，这是不符合事物发展的规律的。科学是一步一个脚印地向前发展，每个人都要依赖前人的工作。当你听说一个突然的、意想不到的发现——仿佛晴天霹雳时，你永远可以确信，它总是由一个人对另一个人的影响所导致的，正是因为有这种相互影响才使科学的进展存在着巨大的可能性。科学家并不依赖于某一个人的思想，而是依赖于千万人的集体智慧，千万人思考着同一个问题，每一个人尽他自己的一份力量，知识的大厦就是这样建造起来的。

这就是当代最伟大的物理学家之一——我甚至认为是最伟大的物理学家——的看法。因此，大家可以理解到为什么科学家往往是国际主义者，为什么科学家现在对科学自由的可能限制极为忧虑。

最后，也许有人会问：“科学家的生活有什么价值呢？”哈代（G.H.Hardy）这样回答这个问题：“（他）给知识增添了一些东西，同时又帮助他人给知识增添了更多的东西；这些东西的价值与伟大的科学家们创造的价值相比，或者与那些身后留下了某种纪念的或大或小的艺术家创造的价值相比，只有程度上的不同，没有性质上的不同。”

第2章　科学的追求及其动机（1985）

正如每个科学家的嗜好、气质以及对科学的态度是千差万别的一样，每个科学家追求科学的动机和涉及的范围也是千差万别的。所以，“科学的追求及其动机”是个很难讨论的题目。除此之外，在科学家的整个一生中，他们的动机很容易发生巨大的变化，的确很难找到一个共同的衡量标准。

我限定自己只反映以前的一些伟大科学家的生活与成就。要反映这些伟人的动机和态度，会受到在交流过程中语义上的严重困难，这种困难体现在：语言上所通行的词和惯用语，有各种不同的评论与判断。实际上，当谈到他人时，人们应该很好地注意屠格涅夫在他的小说《除夕之夜》里，通过英沙诺夫向人们提出的警告：

> 我们正在谈论别人，为什么要牵涉自己？

为了从一个恰当的角度开始我的论述，我将从20世纪20年代中期马约拉纳（Majorana）与费米的谈话开始。当时他俩都只有二十几岁，他们的谈话是一个当时在场的人告诉我的：

> 马约拉纳：每隔500年才有一个类似阿基米德或牛顿这样的科学家出现，而每隔100年就有1至2个类似爱因斯坦和玻尔这样的科学家出现。
>
> 费米：那我将处于一个什么地位呢？
>
> 马约拉纳：理智一点，费米，我并没有谈到你我，我们谈的只是爱因斯坦与玻尔。

I

当谈论到促使人们追求科学的动机时，最好的例子莫过于开普勒。开普勒的独特之处在于他处在一个科学发生了巨大变化的伟大时期，这时科学正在摆脱笼罩在它身上的教条，为日后牛顿的发现铺平了道路。

开普勒不迷信前人的成就，提出了一些包括哥白尼在内的一些伟人所没有提出的问题。开普勒关于行星运动的理论，完全不同于以前所提出的假说；他的关于行星运动的轨道“是椭圆”的断言，更超越了他前人所做的各种各样的改进。在有关行星运动的分析中，开普勒并不偏重于各种几何问题，相反，他提出了以下一些问题：“行星运动的原因是什么？”“如果像哥白尼的假说所指出的那样，太阳是太阳系的中心，那这一事实就应该能够由行星本身的运动和轨道辨别出来。”这些都是物理问题，而不像以前所设想的那样，都是几何构造的问题。

尽管开普勒解决行星运动等问题的方法，完全不同于他以前的任何人，但他的工作仍然是从对观察结果进行仔细分析后得出一般结论的方法，而且是这种方法的一个杰出的例子。他的分析过程漫长并且极其艰辛：他在20多年的时间里，坚持不懈地进行工作，从来没有放弃他的目标。如果用呕心沥血这个词来形容他的努力，也是丝毫不过分的。

开普勒从一开始就认识到，仔细研究火星轨道是研究行星运动的关键，因为火星的运动轨道偏离圆轨道最远，它使得哥白尼的理论显出了严重的缺陷。开普勒还认识到，对第谷·布拉赫准确的观察资料进行分析是整个问题的必不可少的先决条件。开普勒曾经写道：

> 我们应该仔细倾听第谷的意见。他花了35年的时间全心全意地进行观察……我完全信赖他，只有他才能向我解释行星轨道的排列顺序。[1]

> 第谷掌握了最好的观察资料，这就如他掌握了建设一座大厦的物质基础一样。[2]

> 我认为，正当朗高蒙太努斯（Longomontanus）[3] 全神贯注研究火星问题时，我能来到第谷身边，这是“神的意旨”，我这样说是因为仅凭火星就能使我们揭示天体的奥秘，而这奥秘由别的行星是永远揭示不了的……[3]

实际上，开普勒曾千方百计想获得他梦寐以求的第谷的观察资料。如果说他犯了偷窃罪，似乎也并不夸张，因为他自己就曾经承认：“我承认，当第谷死的时候，我正是利用了没有或缺乏继承人这样的有利条件，使第谷的资料由我照管，或许可以说霸占了观察资料。”[4] 他自己

又解释道："争吵的原因在于布拉赫家族有怀疑的天性和恶劣的态度；另一方面，也在于我自己有脾气暴躁和喜欢挖苦人的毛病。必须承认，滕纳格尔（Tengnagel）有充分的理由来怀疑我。我已占有了观察资料并且拒绝把它们交给继承人。"[5]

得到了第谷的观察资料以后，开普勒不断向自己提出了这样的问题："如果太阳确实是行星运动的起源和原因，那么这一事实在行星自身运动中如何体现出来？"他注意到，火星的运动在近日点比在远日点要快些，并且"想起了阿基米德"，于是，他用矢径（连接太阳和火星瞬时位置的矢量）的方法，算出了沿轨道运动的面积。开普勒写道：

> 当我认识到，在运动的轨道上有着无数个点以及相应产生了无数个离太阳的距离，我产生了这样的想法：运动轨道的面积包括了这些距离的和。因为我回忆起阿基米德用同样的方法，将圆面积分解成无数个三角形。"[6]

这就是开普勒于1603年7月发现面积定律的经过。牛顿把它称为开普勒三大定律的第二定律。从此以后，人们都这样称呼面积定律。开普勒用了5年多的时间才建立起这个定律；其实，早在1596年他发表《宇宙的奥秘》这本书之前，他就在探求这一规律，那时他用的方法是把5个规则的多面体与当时已知的6个行星联系起来。

面积定律能够确定轨道上各点的速度的变化，但不能确定轨道的形状。在他得出面积定律的最终表述的前一年，开普勒实际上就摒弃了行星运动轨道是圆的假说。1602年10月他曾写道："行星轨道不是圆。这一结论是显而易见的——有两边朝里面弯，而相对的另两边朝外伸延。这样的曲线形状为卵形。行星的轨道不是圆，而是卵形。"[7]

在做出火星轨道是卵形这一结论之后，开普勒又花了3年时间才确定它的轨道实际上是椭圆，当这一结论确立时，他写道：

> 为什么我要在措词上做文章呢？因为我曾拒绝并抛弃的大自然的真理，重新以另一种可以接受的方式，从后门悄悄地返回。也就是说，我没有考虑以前的方程，而只专注于对椭圆的研究，并确认它是一个完全不同的假说。然而，这两种假设实际上就是同一个，在下一章我将证明这一点。我不断地思考和探求着，直至我几乎发

疯，所有这些对我来说只是为了找出一个合理的解释，为什么行星更偏爱椭圆轨道……噢，我曾经是多么地迟钝啊！[8]

开普勒用了10年多的时间才发现了他的第三定律，即任何两个行星公转周期的平方与它们到太阳的平均距离的立方成正比。1618年，开普勒在他的《宇宙的和谐》一书中表述了这个定律。下面就是开普勒自己对发现这个定律的描述：

> 准确地说，就是在1618年3月8日这天，这一结论显现于我的脑海中。但不幸的是，当我试图用计算来证实它的时候，我又以为它是错误的，因而我抛弃了它。5月15日，这个念头终于又回到了我的脑海中，并且以一种全新的方式使我豁然开朗。它与我17年来对第谷观察资料进行分析所得出的数据吻合得如此之好，以至刚开始的瞬间，我感到我好像在梦幻之中。[9] 。

至此，开普勒呕心沥血的漫长而艰辛的追求，终于结束了。

在他的第一本书《宇宙的奥秘》中，开普勒就说过："但愿我们能够活着看到这两种图像能够相互吻合。"[10] 22年后，当他发现了他的第三定律，从而使得他的梦想得以实现时，开普勒在《宇宙的奥秘》再版中加进了这样的注释："22年后，我们终于活着看到了这一天，并为此感到欢欣鼓舞，至少我是如此；并且我相信梅斯特林（Maestlin）及其他人将分享我的快乐！"[11]

II

布诺特（Max Brod）——一个以出版卡夫卡（Franz Kafka）生前之作而闻名的捷克作者——在他的小说《第谷·布拉赫的赎罪》中，描述和比较了第谷·布拉赫与开普勒的性格特点。尽管布诺特的小说很大程度上与历史事实不符，但他有关开普勒究竟是怎么样的一个科学家的看法，仍是值得摘录的。

> 开普勒使第谷对他充满了敬畏之情。开普勒的全心全意致力于实验工作、完全不理会叽叽喳喳的谄言的宁静心态，在第谷看来，几乎是一种超人的品质。这儿有点不可理喻的地方，即似乎缺乏某种情感，有如极地严寒中的气息……[12]

一个实践的科学家能像布诺特描述的开普勒那样，这么宁静与不露情感？[13]

III

开普勒追求科学的最显著特点是专注于他自己的探求。用雪莱的话来说：“他是一个极为独特的人物。”但是，如果别人也像开普勒那样始终如一，专心致志地追求，他们能否也像开普勒那样成功呢？

首先，以迈克耳孙（Albert Mechelson）为例。在他的一生中，他主要是追求一种测量光速的更准确的方法。他的这种兴趣几乎是出自偶然。当他在美国海军科学院任讲师时，科学院的一位负责人要他准备几篇有关光速的演讲稿。这是1878年的事，它导致1880年迈克耳孙第一次测定了光速。50年后的1931年5月7日，也就是他临死前两天，他着手口述一篇有关他最后观察结果的论文，这篇论文于他死后发表了。由于迈克耳孙的工作，人们对光速的精度从3×10^{-3} 提高到3×10^{-4} ，也就是说提高了10倍。然而到了1973年，其准确性改进到10^{-10} ，这一测量精度今后是不会被超过的。

那么，能够说迈克耳孙50多年的努力都白费了吗？我们暂时撇开这个问题。人们应该注意到在迈克耳孙漫长的工作生涯中，他的许多发明都与他热衷于研究“光波及其应用”有关。例如，因为他研制出了干涉仪，使他首次直接测定了一颗恒星的直径。这是一个惊人的成就。至于在爱因斯坦狭义和广义相对论的公式中起作用的迈克耳孙—莫雷实验，它不可否认地改变了人们对时空观的认识，那更是无人不知、无人不晓的了。奇怪的是迈克耳孙自己，他从来没有对他实验的这一结果表示出一丝快乐。实际上，有记载说，当爱因斯坦于1931年4月拜访迈克耳孙时，迈克耳孙夫人认为有必要私下告诫爱因斯坦：“请不要和他谈以太。”[14]

第二个例子是爱丁顿。在他后半生中，他花了16年多的时间来完善他的“基本理论”。他本人对于他耗费如此巨大的精力有什么看法呢？在他去世的前一年，他曾说道：“过去的16年中，我从来没有怀疑过我的理论的正确性。”[15] 虽然他自己这么说，但他的工作对此后的发展，没有留下任何印迹。

单一的对象和唯一的目标，是科学追求的明智之举吗？

Ⅳ

开普勒的成功事例有力地说明，坚持不懈的努力能导致伟大的且具有开创性的发现。与此同时，很多事例表明许多了不起的想法几乎是自发地产生的。例如，狄拉克曾经写道：“有关泊松括号与电子的相对论性波动方程的研究工作，有些想法突然闪现于我脑海中，我不能够很明确地说清楚这种想法到底是如何出现的，但我感到这种工作是一种‘太容易获得的成功’。”[16]

这一事实，即狄拉克认为他的泊松括号和电子的相对论性波动方程的潜意识的想法是偶然产生的，说明那些有过重大发现的人似乎都想回忆起并珍惜他们取得成就的特殊时刻。狄拉克并非唯一的例子。爱因斯坦就曾这样写道：“1907年，当时我正在写一篇有关狭义相对论综合性论文……突然，我的一生中最得意的想法闪现出来了……即‘对于一个从屋顶自由下落的观测者’（至少在他周围很近的地方）存在着一个非引力场。”[17] 。这种“得意的想法”，我们都知道，后来隐含在他的等效原理之中，这个原理也是他后来的广义相对论的基础。

具有同样往事的例子是费米。我曾经就阿达玛（Hadamard）在《数学领域中的发现心理学》所提出的设想问费米，什么是物理学领域的发现心理。费米回忆了他发现慢中子引起放射性作用的那一时刻。当时他是这样说的：

> 我来告诉你我是如何取得我认为是我一生中最重要的发现。我们当时正努力地研究着中子的感生放射性，但所得到的结果毫无意义。一天，当我来到实验室时，突然冒出了这样的想法，我应该检验检验在入射中子前放一块铅所引起的结果。这一次我特别有耐心，我尽力使铅块安放准确，这与我通常习惯大不相同。很显然，我对某些事情感到不满意，我千方百计找理由拖延在这个位置放上铅块。最后，当我勉勉强强要放上铅块时，我对自己说：“不，我不应该在这个位置放上铅块，我想我应该放上一块石蜡。”事情就是这样发生的，没有任何事先的预告，没有任何有意识的判断。我马上找来一块奇形怪状的石蜡，把它放在本应该是放铅块的位置上。[18]

也许在这方面的一个最动人的故事，是海森伯所讲述的有关量子力学的规律突然闪现在他脑海中的时刻。

……一天晚上，我就要确定能量表中的各项，也就是我们今天所说的能量矩阵，用的是现在人们可能会认为是很笨拙的计算方法。计算出来的第一项与能量守恒原理相当吻合，我很兴奋，尔后我犯了很多的计算错误。终于，当最后一个计算结果出现在我面前时，已是凌晨3点了。所有各项均能满足能量守恒原理，于是，我不再怀疑我所计算的那种量子力学具有数学上的连贯性与一致性。刚开始，我很惊讶。我感到，透过原子现象的外表，我看到了异常美丽的内部结构，当想到大自然如此慷慨地将珍贵的数学结构展现在我眼前时，我几乎陶醉了。我太兴奋了，以致不能入睡。天刚蒙蒙亮，我就来到这个岛的南端，以前我一直向往着在这里爬上一块突出于大海之中的岩石。我现在没有任何困难就攀登上去了，并在等待着太阳的升起。[19]

我们不难理解，在这一特殊时刻海森伯的心情是多么激动啊。我们大家都知道，当时困扰着“旧”的玻尔—索末菲量子理论，有很多困难和佯谬。我们也都知道，长期以来，海森伯与索末菲、玻尔、泡利等人对这些难点和佯谬一直感到困惑不解。当时，海森伯与克拉默斯（Kramers）一起已发表了有关色散理论的论文——在很多方面，这个理论是以后发展的先导。

大约30年后，在他经历了战争的悲伤以及战后的失望、彷徨之后，海森伯提出并阐述了基本粒子理论的思想，对此，我们的反映是什么呢？海森伯夫人在她所写的一本有关她丈夫的一本书中写道：“在一个月光皎洁的夜晚，我们漫步于海因山，他完全被他所看到的幻觉所吸引，一路上，他试图向我解释他的最新发现。他谈到对称性的奇妙之处在于可以把它看作万物的原型；他还谈到和谐、简朴的美以及它的内在的真实性。”[20] 她摘录了一段当时海森伯写给他姐姐的一封信：

实际上，过去的几周对我来说是非常激动的时刻。也许通过类比我能将我所经历的事情很好地描述出来。即在过去的5年中，我费尽心力试图沿一条至今仍不清楚的道路攀登原子理论的顶峰。现在这座顶峰就在我面前，有关原子理论的全部纵横关系，突然地并且清晰地展现于我眼前。这些纵横关系凭借其数学的抽象性以一种

> 简单得不能再简单的方式呈现出来，它是一种我们只能谦恭接受下来的礼物。即便是柏拉图也可能不相信它们是如此之美。因为这些纵横关系不可能被发明，它们从开天辟地以来就一直存在。[21]

你们也许注意到了，大约30年前，在他的有关发现量子力学的基本原理的描述中，他使用了异乎寻常的相似语言和术语。但是，对于他的基本粒子的见解，我们能像海森伯一样，用相同的方式看待吗？对于量子力学，他的思想立即获得承认；相反，他有关粒子物理学的思想，却遭到了反对和驳斥，即使长期与他讨论问题的老朋友泡利也不例外。但是读读海森伯夫人在她的传记的结尾写的一段话，是很令人感动的。

> 带着一种自信的微笑，他曾对我说："我是足够幸运的，当亲爱的上帝还在工作时，我能越过他的肩膀瞧了一下。"对他来说，那就够了，完全够了。这给了他巨大的喜悦和勇气，使他能镇定自如地对付他在这个世上一再遭到的敌意和误解，并且不至于误入歧途。[22]

Ⅴ

汤川秀树（Hidekl　　　　Yukawa）在50多岁时写了一本自传《漫游者》。在这本书中，汤川秀树提出了一个伟大发现对它的发现者所产生影响的另一侧面。人们习惯于这样想：至少从表面上看来，一个很富有而且成果累累的人，以他的一生为素材的自传《漫游者》，应该是对他整个一生的叙述。但是汤川秀树的"漫游"只讲到他伟大发现的论文的发表为止。他以忧郁的笔调写到他那伟大的发现："我再也不想写除此之外的东西了。因为，对我来说，我不懈学习的那些日子是值得留恋的；另一方面，当我想到我的学习时间逐渐被其他事情占据了时，我感到伤心。"[23]

尽管我们所有的人能够分享科学巨匠发现的喜悦，但我们可能对很多很多不太敏锐、运气不佳的人所珍爱和认为值得回忆的东西感到困惑。他们是否像贝克特（Samuel　Backett）的剧作中的人物符拉基米尔和埃斯特拉冈一样，在等待着戈多呢？也许他们会用弥尔顿（Milton）的"就是站着等也是值得的"这句话，来自我安慰？

Ⅵ

我现在讲一讲称赞和同意对一个人追求科学的作用。华兹华斯（Wordsworth）描述牛顿所讲的“一个人独自航行在陌生的智力海洋”情形，并不是我们每个人都能理解的。

我指的是爱丁顿孤独地探求他的基本理论这一事情。尽管他自己对于他的理论的正确性深信不疑，但同时代的人对他的著作漠不关心使他感到很灰心。这种灰心，在他死前几个月给丁戈（Dingle）的一封悲怆的信中表露无遗：

> 我一直试图弄清为什么人们觉得我的理论含糊不清。但我要指出，尽管爱因斯坦的理论被认为是含糊的，但是成百上千的人认为有必要把这种理论解释清楚。我不相信我的理论比狄拉克的理论更深奥晦涩。可对于爱因斯坦和狄拉克，人们却认为克服困难去弄懂这种深奥晦涩是值得的。我相信，当人们认识到必须理解我，并且“解释爱丁顿”成为时尚时，他们会理解我的。[24]

当同时代的人用尖刻的、猛烈的批评表达他们的反对意见时，它能产生一种悲剧性的结果。例如：奥斯特瓦尔德和马赫的强烈批评，使玻尔兹曼深感失望，以致他最后自杀了。他的孙子费莱姆（Flamm）曾写道：“他是他的思想的殉道者。”又如，点集合和无穷阶等现代理论的创始人乔治·康托（Georg Cantor），因为他的老师利奥波德·克罗内克（Leopeld Kronecker）痛恨并敌视他以及他的理论，结果康托精神失常。事实上，他的晚年是在精神病院度过的。

Ⅶ

一种完全不同于以上所讲述的例子是卢瑟福。

我们先回想一下他的经历。1897年，他引入了一种新的命名方式，将放射性辐射分解成α粒子、β射线和γ射线。1902年，他将放射性衰变的规律公式化——在物理学中首次用概率和不确定性等术语来表达物理公式，而且也是量子力学概率表述的先驱，这种概率表述在25年后已变得非常普遍。1905年至1907年间，他和索迪一起，系统地阐述了放射性衰变位移的规律，并确定了α粒子就是氦核；和波特伍德一起，他创立了用放射性来确定岩石和矿物的年代的方法。1909年至1910年，这一时期出现了盖革和马斯顿实验，这一实验发现了α粒子的大角散射，接着

卢瑟福系统地阐述了散射规律并提出了原子构造的有核模型。接着，在1917年，他又首次在实验室中完成了原子嬗变的实验：即用α射线轰击N^{14}，使之变成O^{17}和一个质子。20世纪20年代，他明确了α射线和γ射线光谱间的相互关系。1932年，福勒称之为“创造奇迹的一年”，考克饶夫和瓦尔顿发现了用人工衰变的方法将Li^{7}转变为两个α粒子，布莱克特在宇宙射线表中发现了正电子，以及查德威克发现了中子。创造这些“奇迹”的人都在剑桥的卢瑟福的卡文迪许实验室中工作。1933年，卢瑟福和奥利芬特一起发现了H^{3}和He^{3}。

卢瑟福对他自己的发现的态度，在一次答复别人的话中，表示得很清楚。有一次，当他有一项重大发现时，别人说道：“卢瑟福，你总是处在事业的波峰之上。”卢瑟福答复说：“我自己创造了这个波，难道不是吗？”不知道为什么，从卢瑟福的有利条件来看，似乎他所说的任何话都是正确的，甚至包括下面一句话，有人问他，他是否鼓励他的学生学习相对论时，他回答说：“我不会让我的孩子们浪费他们的时间。”如果世上真有快乐的战士的话，那么卢瑟福就是其中的一个。

Ⅷ

到现在为止，我已通过列举一些科学巨匠的生活中的事件，来描述他们对科学追求的各个侧面。现在我回过来讲一些更一般的例子。同样，我从一个实例开始。在迈克耳孙晚年，有人问他，为什么他花这么多时间去测量光速。据说他的回答是：“它太有趣了。”不可否认，在对科学的追求中，“有趣”起了一定的作用，但是“有趣”这个词并不意味着缺乏严肃。确实，在《简明牛津字典》中，“有趣”（fun）被解释成“滑稽举动”（drollery）。但是我们能够肯定，当迈克耳孙把他生活的主要兴趣描述成“有趣”时，绝对没有那样的含义。然而，我们赋予迈克耳孙所用的“有趣”这个词的准确含义又是什么呢？更一般地说，愉悦和乐趣的作用是什么呢？

尽管愉悦和乐趣经常被用来描述一个人在科学中的奋斗结果，但失败、挫折及绝望起码也是追求科学的人常常碰到的事。毫无疑问，克服了困难，将会给人带来最终成功的快乐。那么，失败纯粹是科学追求中一个否定的方面吗？

在20世纪20年代中期和末期，随着量子力学的基本原理的建立，物

理学有了迅猛的发展。狄拉克对这种发展的描述是：

> 将它描述成一种游戏（game），一种非常有趣的游戏，这种说法是很恰当的。无论什么时候，某人解决了一个小问题，他都能就此写一篇论文。当时，对于一个二流的物理学家来说，很容易做出一流的工作。此后，再也没有一个如此辉煌的时代了。[25]

J. J.汤姆逊在威斯特敏斯大教堂为纪念瑞利勋爵所做的讲话，是颇值得注意的。他说：

> 有一些科学巨匠的魅力在于他们对一个课题首先做出了说明，在于他们提出了以后被证明是很有成效的新思想。还有一些科学巨匠的魅力也许在于他们完善了某一课题，使它变得具有连贯性和明确性。我认为瑞利勋爵属于第二类。[26]

J. J.汤姆逊的这种评论有时被认为是一语双关。但是人们可以得出这样的结论：瑞利在气质上更愿意致力于艰难的问题，而不满足于做狄拉克所说的“第二流的物理学家能做出第一流的工作”的物理学“辉煌时代”的游戏。

上述有关瑞利气质的结论，又引出了另一个问题：当一个科学家已取得名望之后，他继续追求科学的原因是什么呢？这些原因在多大程度上带有个人的动机呢？在多大程度上与美学标准（如秩序和图案、形式和内容的直觉感受）有关联呢？美学标准和个人动机是相互排斥的吗？责任感起作用吗？我所说的责任不是通常意义上所说的某人对他的学生、同事及他的社会所承担的责任，而是说对科学本身的责任。那么，为科学而追求科学时，责任的真正含义又是什么呢？

让我最后再讲讲另一侧面。哈代在结束《一个数学家的自白》这篇文章时，用了下面这样一段话：

> 我一生的经历，或者说与我有相同经历的任何数学家的一生经历是：我给知识增添了一些东西，同时又帮助他人给知识增添了更多的东西；这些东西的价值与伟大数学家们的创造性价值相比，或者与那些在身后留下了某种纪念的或大或小的艺术家们的创造性价值相比，只有程度上的不同，没有性质上的不同。[27]

哈代的叙述是针对数学家而言的，但它同样适于所有的科学家。我希望你们特别注意他所指的给后人留下一些纪念，正是后人能够据以做出判断的东西。那么，后人的判断——科学家本人永远也不会知道的东西——在多大程度上也是科学家追求科学的一种有意识的动机呢？

IX

对科学的追求常被比喻成攀登一座很高但又不是高不可攀的山峰。我们当中有谁能够奢望（即使是在想象中），在一个晴朗无风的日子里去攀登珠穆朗玛峰并达到它的顶点，在宁静的空中，纵览那在雪中白得耀眼的一望无涯的喜马拉雅山脉呢？任何人都不可能奢望看到自然界和宇宙的上述情景。但是，仅仅站在峡谷的底端等待着太阳越过干城章嘉峰，也未免太俗不可耐了。

注释

[1]Letter to Masetlin，16-26 February 1599，in Johannes Kepler gesammelteWerke, ed. w.von Dyck and M.Caspar（Munich，1938），13：289（hereinafter cited as GesammelteWerke）.Quoted in Arthur Koestler, The Sleepwalkers（London：Hutchinson，1959），p.278.

[2]Letter to Herwart，12 July 1600，Gesammelte Werke 14：218（Koestler, p. 104）.

[3]Kepler, Astronomia nova, in Gesammelte Werke, vol. 3，dedication（Koestler, p.325）.

[4]Letter to Heyden, October 1605，Gesammelte Werke，15：231（Koestler, p. 345）.

[5]Letter to D. Fabricius，1 October 1602，GesammelteWerke，15：17（Koestler, p.345）.

[6]Kepler, Astronomia nova, chap. 40（Koestler, p.327）.

[7]Kepler, Astronomia nova, chap. 44（Koestler, p.329）.

[8]Gesammelte Werke，15：314（Koestler, p. 333）.

[9]Gesammelte Werke，16：373（Koestler, pp. 394-395）.

[10]Gesammelte Werke, vol. 1，chap.21（Koestler, p.260）.

[11]Ibid.，note 7.

[12]Max Brod, The Redemption of Tycho Brahe（New York：Knopf，1928），p. 157.

[13]When he wrote The Redemption of Tycho Brahe, Max Brod was a member of the small circle in Prague that included Einstein and Franz Kafka. Brod's portrayal of Kepler is said to have been influenced by his association with Einstein.Thus, Walter Nernst is reported to have said to Einstein：“You are this man Kepler.”see Philip Frank, Einstein：His Life and Times（New York：Knopf，1947），p.85.

[14]Dorothy Michelson Livingston, The Master of Light：A Biography of Albert A. Michelson（Chicago：University of Chicago Press，1974），p.334.

[15]A. Eddington, Dublin Institute of Advanced Studies A，1943，p.1.

[16]P. A.M.Dirac，“Recollections of an Exciting Era.”inHistoryof TwentiethCenturyPhysics, Proceedings of the International School of Physics，“Enrico Fermi”（New York：Academic Press，1977），pp.137-138.

[17]A. Pais, SubtleistheLord（New York：Oxford University Press，1982），p.178.

[18]S. Chandrasekhar, Enrico Fermi：Collected Papers，2 vols.（Chicago：University of Chicago Press，1962），2：926-927.

[19]W. Heisenberg, Physics and Beyond：Encounters and Conversations（New York：Harper&Row，1971），p.61.

[20]E. Heisenberg, InnerExile, trans.S.Cappellari and C.Morris（Boston：Birk H user，1984），p.143.

[21]Ibid.，pp. 143—144.

[22]Ibid.，p. 157.

[23]H. Yukawa, Tabibito（The Traveler）（Singapore：World Scientific Publishing，1982），p.207.

[24]J. G.Crowther, British Scientists of the Twentieth Century（London：Routledge&Kegan Paul，1952），p.194.

[25]P. A.M.Dirac, DirectionsinPhysics（New York：Wiley，1978），p.7.

[26]R. J.Strutt，4th Baron Rayleigh, LifeofJohnWIlliamStrutt, Third BaronRayleigh, O.M.，F.R.S.（Madison：University of Wisconsin Press，1968），p.310.

[27]G. H.Hardy, AMathematician'sApology（Cambridge：Cambridge University Press，1967），p.151.

第3章　诺拉和爱德华·赖森讲座　莎士比亚、牛顿和贝多芬：不同的创造模式（1975）

对弥尔顿（Milton）的贬抑性批评，艾略特（T.S.Eliot）曾这样说，“只有他那个时代最有才能的诗人的裁决”才能令他折服。十年以后，也许是因为情绪比较好，他补充道：“在文学评论领域内，学者和作家应该相互取长补短。如果作家多少有点学问，他的批评肯定会更好；如果学者多少能体验到遣词造句之困难，那么他的批评也肯定会更妙。”按照同样的标准，任何人如果敢于大胆地探索艺术工作者和科学工作者的不同创造模式，那么他一定是一个艺术领域及科学领域的实干家，同时也一定是一个学者。若仅仅是科学领域或艺术领域的工作者，那是不够的。我是一个在自然科学的某个小巷里独自游荡的流浪汉，对周围世界知之不多，但竟然向自己提出了一个包括艺术和科学的问题，这肯定使我力不从心，因此，首先我请求你们要有耐心。

考虑到各人的兴趣、气质、悟性的千差万别，我们会问：我们能够辨别艺术工作者和科学工作者创造模式的主要差异吗？我想，这个问题可以用下述方法探讨：首先，考察莎士比亚、牛顿和贝多芬的创造模式，他们在各自的领域内都达到了人类成就的顶峰，这是举世公认的，然后从这些成就辉煌的创造模式的异同中，寻求对个别情况行之有效的更普遍的结论。

Ⅰ

我先从莎士比亚谈起。

同伊丽莎白女王一世时代所有人受的教育一样，莎士比亚受的教育很少。虽然这些教育使莎士比亚感到心满意足，但是莎士比亚从来没有信服过他学到的知识。下面两段话清楚地表明了这一点：

> 死啃书本，终无所获，
>
> 引经据典，吓唬别人。

或：

> 噢，学问，鬼才知道是什么玩意！

虽然如此，1587年，当23岁的莎士比亚到达伦敦时，他没有像洛奇、基德那样有利的社会背景；也没有皮尔、李里、格林、马洛、纳什[4] 在牛津或剑桥大学镀金多年的优势。毋庸置疑，莎士比亚敏锐地意识到了自己的短处和不足，于是他只有通过各种渠道阅读和吸收知识来弥补自己的不足。霍林斯赫德（Holinshed）的《英格兰、苏格兰、爱尔兰历史》第二次修订本的出版，对莎士比亚来说正是时候，这鼓舞他去创作即将问世的历史剧。

到1592年时，莎士比亚已写成了《亨利四世》的三个部分和他的早期喜剧：《错误的喜剧》《爱的徒劳》和《维洛那二绅士》。那年，这些剧本的成功遭到了罗伯特（Robert）、格林的恶毒攻击。格林比莎士比亚大6岁，那时他已是伦敦文学领域内几个显赫的人物之一。格林的攻击之作是他死后才出版的。他死得太早了一点，死因是由于一次致命的晚宴，据说是由于“莱茵酒和腌鲱鱼”。死后他的文章是他“留下的一颗定时炸弹”。攻击之词中有这样一段：

> 有一只突然飞黄腾达的乌鸦，用我们的羽毛装饰打扮起来，他的“演员的外表下包藏着一颗虎狼之心”，他自认为像你们中的佼佼者一样，可以挥笔写出无韵诗。他是个地道的“打杂儿”，却妄自认为是国内独一无二的“全才”。

格林的攻击清楚地表明，他把莎士比亚看成一个局外人、一个入侵者。莎士比亚没有大学学历，因此他不属于贵族圈子里的人。

尽管早期取得了一些成功，但对莎士比亚来说，他仍然只是一个演员和一个写剧本的人而已，又因为瘟疫时常发生，伦敦的剧院也因此经常关闭，生活充满着不安定。但从1590年起，情况大为改观，莎士比亚找到了一个庇护人，一个朋友，还得到了爱。

莎士比亚的庇护人是年轻的南安普敦伯爵（Earl of Southampton），他在1591年才成年。随后的4年中，莎士比亚对南安普敦的强烈感情，对于他的艺术发展和为他所展现的机会是有决定性意义的。莎士比亚的天才发挥得淋漓尽致，爆发出空前的创造力。除上面提

到的剧本外，他还写了《威尼斯商人》《驯悍记》和《理查三世》。献给南安普敦伯爵的两首精采的叙事诗《维纳斯与阿都尼》《鲁克丽丝受辱记》，也是这一时期的作品。

在1592年到1595年间，莎士比亚写了许多十四行诗，作为对南安普敦伯爵保护之恩的报答。《十四行诗》比他的其他作品更具有自传性质，它们使我们得以明白莎士比亚对自己和对他的艺术的态度；同时也可以让人们看出，他依赖南安普敦的友谊和庇护的程度。

莎士比亚和南安普敦友谊的历程并不平坦，这是因为他们的年龄、地位存在着差异——一个是诗人，一个是贵族庇护人；此外，还有莎士比亚的情人——十四行诗中的“黑女人”引起的纠葛，她撇开了莎士比亚，转向了容易动情的伯爵。在十四行诗中，莎士比亚倾泻了他强烈而真挚的感情：

当我受尽命运和人们的白眼，

暗暗地哀悼自己的身世飘零，

徒用呼吁去干扰聋聩的昊天，

顾盼着身影，诅咒自己的生辰。（29）[5]

为抵抗那一天，要是终有那一天，

当我看见你对我的缺点蹙额，

当你的爱已花完最后一文钱，

被周详的顾虑催去清算账目；

为抵抗那一天，当你像生客走过，

不用那太阳——你眼睛——向我致候，

当爱情已改变了面目，

要搜罗种种必须决绝的庄重的理由；

为抵抗那一天我就躲在这里，

在对自己的恰当评价内安身，

并且高举我这只手当众宣誓，

为你的种种合法的理由保证：

抛弃可怜的我，你有法律保障，

既然为什么爱，我无理由可讲。（49）

他们的关系，至少在莎士比亚看来，已到了非常脆弱的程度，莎士比亚甚至想到了死：

我死去的时候别再为我悲哀，

当你听见那沉重凄惨的葬钟，

普告给全世界说我已经离开，

这龌龊世界上去伴最龌龊的虫；（71）

莎士比亚还感到，如果失去了南安普敦的爱，他的生命也将无法存在，生命与友谊同生死、共存亡。

但尽管你不顾一切偷偷溜走，

直到生命终点你还是属于我。

生命也不会比你的爱更长久，

因为生命只靠你的爱才能活。

因此，我就不用怕最大的灾害，

既然最小的已足置我于死地。

我瞥见一个对我更幸福的境界，

它不会随着你的爱憎而转移；

你的反复再也不能使我颓丧，

既然你一反脸我生命便完毕。

哦，我找到了多么幸福的保障：

幸福地享受你的爱，幸福地死去！

但人间哪有不怕玷污的美满？

你可以变心肠，同时对我隐瞒。（92）

尽管这种含糊的话在十四行诗里经常出现，但莎士比亚对前途充满信心的话，也时而在他的诗中喷薄而出。在著名的第55首十四行诗中，他的激情流露得酣畅感人：

没有云石或王公们金的墓碑，

能够和我这些强劲的诗比寿；

你将永远闪耀于这些诗篇里，

远胜于那被时光涂脏的石头。

当这残暴的战争把铜像推翻，

或内讧把城池荡成一片废墟，

无论战神的剑或战争的烈焰，

都毁灭不了你的遗芳的活历史。

同时，南安普敦的另一个被庇护人马洛，作为一个危险的敌手出现了。为了抵消莎士比亚的《维纳斯与阿都尼》一诗的影响，马洛写下了他的诗作《希罗和李安德》。莎士比亚承认了马洛的优势，并对这一敌

手表示出了某种不安：

哦，我写到你时多么气馁，

得知有更大的天才利用你的名字，

他不惜费尽力气去把你赞美，

使我钳口结舌，一提起你的声誉！

但你的价值，像海洋一样无边，

不管轻舟或艨艟同样能载起，

我这莽撞的艇，尽管小得可怜。

也向你茫茫海心大胆行驶。

你最浅的滩濑已足使我浮泛，

而他岸然驶向你万顷汪洋；

或者，万一覆没，我只是片轻帆，

他却是结构雄伟，气宇轩昂：

如果他安全到达，而我遭失败，

最不幸的是：毁我的是我的爱。（80）

1593年马洛在一次不幸的殴斗中死去，莎士比亚曾在《皆大欢喜》中借试金石的口说：

要是一个人的诗不被人懂，他的才情也得不到应有赏识，那比小客栈里开出一张大账单来还要命。

在同一剧本中，莎士比亚给马洛做了不寻常的赞颂，称他为“谢世的牧羊人”，还引用了马洛的诗句：

哪个情人不是一见钟情？

不久，莎士比亚与“黑女子”的不幸插曲也结束了：

最反复是我；

我对你的一切盟誓都只是滥用，

因而对于你已经失去了信约。（152）

在有关南安普敦组诗的最后一首十四行诗中，莎士比亚表现了胜利者的喜悦：

不，请让我在你的心里长保忠贞，

收下这份菲薄但由衷的献礼，

它不掺杂次品，也不包藏机心，

而只是你我间相互致送诚意。（125）

是的！“菲薄但由衷”，“不掺杂次品”和“只是你我间”。

1594年，南安普敦伯爵给了莎士比亚大约100英镑，作为张伯伦勋爵公司创建的一份股金。由于未来有了保证，莎士比亚原有的勇气被激发出来，天才也成熟了。这一年，他写出了《仲夏夜之梦》，这是他的第一部伟大杰作。不久《罗密欧与朱丽叶》《皆大欢喜》《无事生非》相继写成。接着，莎士比亚再次转向历史剧，写出了《约翰王》《亨利四世》上、下篇和《亨利五世》。在所有这些历史剧中，只有一个英雄，那就是英格兰；而且在这些剧本中莎士比亚生动表现了“那个时代的真实面貌”。

许多人认为《亨利四世》上、下篇是莎士比亚历史剧的顶峰。由于福斯塔夫（Falstaff）[6]的形象深深印入了人们的脑海中，它们理所当然地是最好的戏剧。有人说：“福斯塔夫进入英国文学的同时，堂吉诃德进入了西班牙文学，但他们进入的方式截然不同。”

莎士比亚成果最大的是他的“中年时期”，它开始于《仲夏夜之

梦》，结束于《汉姆雷特》（1600—1601）。

在《汉姆雷特》中，莎士比亚表达了他的戏剧思想，也表达了他对正在兴起的敌手本·琼生[7]以及黑僧剧院（Blackfrjar’s theater）重视机智和时尚戏剧的反应。我们发现汉姆雷特在对演员的指令中（剧中之剧）这样说：[8]

> 任何过分的表演都是和戏剧的原意相悖的，自有戏剧以来，戏剧的目的始终是反映自然，显示善恶的本来面目，给它的时代看一看它自己演变发展的模型。

莎士比亚在这里主张“时代的风貌”能通过戏剧表现出来——的确他已在历史剧中表现了他自己的时代。

下面一段话也许有告诫本·琼生和“革新者”的暗示：

> 啊！我顶不愿意听见一个披着满头假发的家伙在台上乱嚷乱叫，把一段感情片片撕碎，让那些只爱热闹的低级观众听了出神，他们中间的大部分是除了欣赏一些莫名其妙的手势以外，什么都不懂……

> 啊！我曾经看见几个伶人演戏，而且也听见有人为他们极口捧场，说一句比喻不伦的话，他们既不会说基督徒的语言，又不会学着基督徒、异教徒或者一般人的样子走路，瞧他们在台上大摇大摆，使劲叫喊的样子，我心里就想一定是什么造化的雇工把他们造了下来：造得这样拙劣，以至全然失去了人类的面目……

> 啊！你们必须彻底纠正这一弊病。

在继《汉姆雷特》后的两个剧本《终成眷属》和《一报还一报》中，有证据表明，那时莎士比亚的精神已处于崩溃的边缘。他不再对人和事物抱有幻想——也许这时的心情最适宜写伟大的悲剧。正如以研究伊丽莎白一世和莎士比亚而著名的学者罗斯（A.L.Rowse）所说，伟大的悲剧“显示了他精神极度的紧张和身心交瘁”；他还写道：

> 像所有的有意义的工作一样，我们的研究只能放在一些重点上，要么专门研究文学的一面，要么其他一些个人问题……如果莎

> 士比亚想和他的敌手本·琼生比个高低，那他就必须在悲剧上比试比试。在创作悲剧这方面，他做出了最可贵的努力，他的才能施展到最大限度，这样，作为一个作家他完成了他的使命……有足够的证据表明：作为一个作家他并不是不在乎他的名誉和成就，他的雄心远不止于此。于是问题转了一个圈子后只能认为，他的工作实际是出于他个人的考虑。

当莎士比亚作品完成后，本·琼生只能将他与伟大的悲剧作家埃斯奇勒斯、沙孚克里斯和尤里皮蒂[9]相比较。

1604—1608年，《奥赛罗》《李尔王》《麦克白》《安东尼与克里奥佩特拉》和《科利奥兰纳斯》相继问世。这些伟大的剧本简直让人感到震惊，它们彼此之间完全不同，没有持续的灵感，是不可能接连写出这么多伟大剧作的。

黑兹里特[10]对这些悲剧的概括是：

> 《麦克白》《李尔王》《奥赛罗》和《汉姆雷特》通常被认为是莎士比亚的四部主要悲剧著作。《李尔王》表现了深厚强烈的感情；《麦克白》表现了行动的敏捷和想象的奔放；《奥赛罗》表现了情感的渐进和急剧变更；《汉姆雷特》表现了思想和感情微妙的发展。如果说天才的力量在这些剧本中的每一个剧本里都表现出来，是一件令人惊诧的事情的话，那么它们的变化多端也绝不逊色。它们好像是出自同一头脑中的不同创造，它们之间找不到一丝一毫的关联。这种特征和创造性，的确是真理和天性的必然结果。

黑兹里特没有把《安东尼与克里奥佩特拉》算在伟大的悲剧之中，但今天许多人认为它同样伟大。艾略特曾对《安东尼与克莉奥佩特拉》做过极为敏感的分析，他说：

> 这是一部为成熟演员和成熟观众写的戏剧，乳臭未干的少男少女们，不管他们是演员还是观众，都不能领悟到这些中年恋人的感情……《安东尼与克里奥佩特拉》的成功之处就在于，在生活的同一侧面将英雄和可怜虫有机地融合在一起。马洛似乎也能使他的人物形象同样高贵。德莱顿[11]后期的剧本中，其主题也几乎是同样的。但只有莎士比亚能使他们不但高贵而且还有人类的软弱，没

有人类的软弱就没有悲剧的可怕和伟大，而究其原因，就在于莎士比亚学会了用诗的语言去表达事物，而其他人甚至用散文也表达不了它。

有人认为，继伟大的悲剧之后创作的剧本《雅典的泰门》《泰尔亲王配力克里斯》和《辛白林》都显示出了心情紧张后的疲劳。正如A.L.罗斯评论说："这些年似乎有一个间歇停顿，看来真是如此。"但是艾略特表示了相反的意见：

> 越是后面的戏，写起来难度越大。在谈到和听到《安东尼与克里奥佩特拉》时，我们的惊诧程度在许多地方可以用下面的话表达出来："我从来没想到那可以用诗的语言表达出来。"对后面的几个剧本，我们在惊奇的时刻，可以恰当地这样说："我们从来没有想到那样的事能完全表达出来。"最后的几个剧本，我指的是《辛白林》《冬天里的故事》《泰尔亲王配力克里斯》和《暴风雨》。在这些剧本中，莎士比亚为了给我们展现更深邃的感情世界，已经放弃了平常的现实主义……

无论怎样，莎士比亚的最后三部剧本《冬天里的故事》《暴风雨》和《亨利八世》，更易于理解。至少，莎士比亚本性所具有的自信、沉着得到了进一步的证实。《冬天里的故事》是一部非常美丽和感人肺腑的戏剧。黑兹里特称它为"我们戏剧作家写出的最杰出的剧本之一"。同时，研究莎士比亚的某著名学者写道："《冬天里的故事》是无可挑剔的，甚至无法用语言称赞它。"

在最后第二部剧本中，莎士比亚曾设法寻找一些新鲜的东西，以解决一个深奥的主题，这一主题至今仍困惑着我们，那就是他塑造的凯利班[12]这一形象。可以说他为我们具体描述了当今的一个中心问题。但是《暴风雨》中所表现的情绪，已一去不复返了。

> 我们的狂欢已经终止。这些演员，
>
> 我曾经告诉过你，原是一群精灵；
>
> 他们都已化成烟雾而消散了：
>
> 如同这虚无缥缈的幻景一样，

入云的楼阁，瑰伟的宫殿，

庄严的庙堂，甚至地球自身，

所有这一切都将同样消散，就像这一场幻景，

连一点烟云的影子都没有留下。

莎士比亚的最后一部剧本终于又回到以英格兰为背景的历史剧中去了。他的历史剧，从《亨利五世》和《理查三世》开始，最后以《亨利八世》和伊丽莎白的出世告终。大主教坎特布里总结性的演讲一开始就这样祈祷：

这位皇室的公主——愿上帝永远在她周围保护她——

虽然还在襁褓，已经可以看出，

会给这片国土带来无穷的幸福……

这是伊丽莎白时代到来的预言。它提供给莎士比亚一个赞颂女王的极好机会。在女王1603年逝世的时候，他没有献上他的颂词，现在有机会总结伊丽莎白时代的特征了。正如A.L.罗斯在他的《莎士比亚传》的结尾写的那样：

这也是莎士比亚的结局，像一条色彩斑斓的巨蛇蜷曲着身子，闪闪发光——象征智慧与永生——他的著作等身，完美超群。

本·琼生的颂词曾预言：

他不属于一个时代，而是属于一切时代。

这里我援引两位当代作家的话，作为本小节的结束语。

弗吉尼亚·伍尔芙[13]冥思苦想一番后，怎么也想不出莎士比亚是如何“遣词造句”的，她在她的日记中写道：

以我们对文学的理解来评断，我可以说莎士比亚完全超越了文学。

艾略特把莎士比亚概括为：

> 莎士比亚的准则就是从始到终持续不断地发展，在每一部戏剧中，戏剧的情节和诗一般的技巧的发展，似乎越来越由莎士比亚的感情状态支配，而其感情状态又由当时情感成熟的具体情况决定……我们可以毫不犹豫地说，他的每部剧本的全部意义并不仅限于它自身，要了解某一剧本的全部意义，必须知道该剧本是什么时候写的，它和莎士比亚其他剧本以及前后剧本的关系如何。我们要想了解莎士比亚的任何一本剧作，就必须知道他的全部著作。当时没有一个戏剧作家能达到这样完美的境界……

莎士比亚事业的发展是如此令人惊异，像《汉姆雷特》一样，它可以触动绝大部分人的心灵最深处的情感，并诱发出无比丰富的想象。把莎士比亚的工作和某一自然规律相对应的话，它就像彗星向地球靠近然后又渐渐地离开。莎士比亚也是渐渐地离开人们的视野，直到消失在他个人的神秘世界之中。

Ⅱ

现在，我以更不安的心情来谈论贝多芬。由于我在音乐方面没有什么造诣，谈论他会更感到吃力。

贝多芬1792年到维也纳时，已有22岁，他当时一定非常谨慎。他拜海顿（Haydn）、申克（Schenk）、阿尔布雷希特贝格尔（Albrechtsberger）和萨利埃瑞（Salieri）为师。我们可以猜测，当初他是想从他们身上学到一些东西。他清楚地注意到他从他们那儿学到的东西，并不能改变他自己的音乐思想，因而，一旦他发现他在钢琴上即兴演奏作品的高超技巧能胜过维也纳的每一位音乐家时，他就忍不住了，甚至有时还表现出一种挑战的意味。这样，当海顿轻视他的三个三重奏中的第一个三重奏作品1号时，贝多芬认为这恰恰证实了自己的看法，即它是三个中最好的一个，海顿的轻视是由于嫉妒和怨恨。

这时，贝多芬渴望得到伟大的声誉，他似乎毫不怀疑自己的超群才能足使他免遭所有不幸。他的这一态度在他给冯·策斯卡尔（Von Zmeskall）的信中表现得十分清楚。

> 见鬼去吧！我对你的整个道德体系不屑一顾。能力就是出类拔萃者的道德，这也是我的道德。

这种极端的自信，就来自于这种力量道德观，它注定使他遭受最大的痛苦和磨难。

当他28岁时，第一次耳聋的症状出现了。对听力减退最初的反应是感到痛苦，不时显得暴躁不安。3年以后，他写信给阿芒达牧师（Karl Amenda）说：

> 你的贝多芬遭到了非常的不幸，和大自然的造物主发生了争吵。我常常诅咒造物主，他常常毫无缘由地将他创造的东西遗弃，以致最美丽的花蕾因此常常被糟踏、凋谢了。你只要想一想，我最高贵的部分，我的听觉，大大地衰退了，这是多么可怕的事！

但他的意志并没有消沉，他继续说道：

> 我决心扫除一切障碍……我相信命运不会抛弃我，我恐怕需要充分估量自己的力量……我将扼住命运的喉咙。

我们从他在1802年写的著名的埃林耿希泰脱（Heiligenstadt）[14] 遗嘱中，可以很好地了解当时贝多芬的精神状态。这个遗嘱在他死后，才在他的手稿中发现。遗嘱是如此坦诚，我真想写出它的全部内容，但下面的一段已足够说明一切：

> 每当我旁边的人听到远处的笛声而我听不见时，或他们听见牧童歌唱而我一无所闻时，真是何等地屈辱！这种体验几乎使我完全陷于绝望：我差一点想结束自己的生命——是艺术，仅仅是艺术把我从死亡线上唤回。啊！在我尚未把我感到需要谱写的每一乐章完成之前，我觉得不能离开这个世界。

贝多芬承认他曾打算自杀，正是他未完成的艺术这一力量挽救了他，这种力量在20年后得到了反响：

> 我作为一个普通的人，仅仅为我的艺术和未完成的职责而活着。

很显然，贝多芬早期力量道德伦理观，随着他的耳聋而土崩瓦解了。但是，它像一只长生鸟又再生了，只有靠着它才能使他的创造力得以实现。这样，到1807年他又写出了他的第三首《拉苏莫斯基》弦乐四重奏。他似乎已经完全从折磨中恢复过来，我们在作品旁的空白处看到：

即使对于艺术，也不必再对它掩饰你的耳聋……

大家都认为，和命运做斗争的宏伟场面，在他的《第五交响曲》里表现得最淋漓尽致。

这段“中年期”的高强度创作大约持续了10年。到40岁出头时，贝多芬已谱写了8首交响曲，5首钢琴协奏曲，1首小提琴协奏曲，25首钢琴奏鸣曲，11首四重奏曲，7首序曲，1部歌剧，1首弥撒曲。贝多芬在取得辉煌成就后，从42岁起有7年没有创作。这一定是他在沉思、反省。继沉寂时期而来的成果，也许在音乐史上是绝无仅有的。

从1801年的《第一交响曲》到1812年的《第八交响曲》，在本质上是同一个贝多芬，一个常人所能理解的贝多芬。但是贝多芬的《第九交响曲》、《D大调庄严弥撒》、最后4首钢琴奏鸣曲，尤其是最后5首四重奏，所有这些则完全是另一个贝多芬。贝多芬的学生，切尔涅（Czerny）就不能理解他这一时期的音乐，他试图把它归结于贝多芬的耳聋。

贝多芬的第三种风格起始于他逐渐耳聋的时期……

这导致他最后三部钢琴奏鸣曲的独特风格……出现了许多不协调的和音……

根据各方面综合考虑，贝多芬最后几部四重奏是他成就的“珠穆朗玛峰”，下面的说法再典型不过了：

它们是无与伦比的。

它们是无法用语言描述和分析的。

最后几首四重奏是独一无二的，对贝多芬来说是独一无二的，

在所有的音乐中也是独一无二的。

但一定有许多人会这样说：没有人能说出这些四重奏的真正含意，我们仅能确信的是，它们表达的思想境界在其他任何地方都不能找到。用描述牛顿思想的名言，“有如独自穿过陌生的思想海洋”来描述最后时期的贝多芬的思想，再恰当不过了。

F大调四重奏第十六号是贝多芬最后完成的作品，它为贝多芬伟大的一生提供了一个辉煌的结尾。对于这首四重奏，J.W.N.沙利文这样评价道：

> 这是一个极度宁静的人所创造的作品，这是一个曾搏击长空但如今一切已成往事的人所拥有的宁静。这一特点最充分地显示在他最后一个乐章的主题句上：“一定是这样吗？一定是！”

回顾贝多芬的生活和创作，沙利文这样概括道：

> 要想了解贝多芬，最有意义的事实之一就是他的工作自始至终都是在有机地发展着……贝多芬创作的最伟大的乐曲是最后几首四重奏，从后往前看，每十年，他的音乐都较前十年有更大的进步。

这一概括和我们前面援引的艾略特对莎士比亚的概括有异常相似之处。莎士比亚、贝多芬两人早年克服生活危机的方式，他们不断成熟的思想，他们的创作和全部生活的有机结合，他们生命后期的伟大杰作，甚至在《暴风雨》和四重奏16号中显示出的告别心情，所有这些，的确有惊人的相似之处。

III

现在来谈牛顿。

伊萨克·牛顿，一个遗腹子，于1642年圣诞节那天降临人世，凯因斯（M.Keynes）曾贴切地写道：“这是最后一个奇婴，东方圣人也得向他致以真诚而恰如其分的敬意。”[15]

牛顿生活中最引人注目的事件就是他突然爆发出的天才。儿时的他不是一个神童，当他1661年去伦敦剑桥的时候，可能除了基本的算术之

外，他知道得很少。不要忘记，那时与伽利略、开普勒、笛卡儿等人名字有关联的科学思想的新轮廓，还没有在剑桥和牛津引起人们的重视。然而，到1664年牛顿22岁时，他的天才似乎已经含苞欲放了。牛顿晚年回忆说，他“在这个时候（1664—1665）发现了无穷级数方法”。事实上，牛顿写出了一些摘记，后来成为一篇连贯的论文，题目是“无穷多项式的分析”，并答应巴罗把它送给柯林斯（Collins）。并约定，无论如何都不署他的名。这个约定后来收回了，但我们在这儿第一次看到了牛顿的一种秉性，这一秉性伴随牛顿终生。

1665年夏，由于瘟疫流行，剑桥大学停课，许多人被疏散出城。牛顿回到伍尔兹索普，这时他的天才像鲜花一般盛开。在科学思想史上，这是无可超越的一段时期。但是直到许多年以后，整个世界才知道牛顿在伍尔兹索普的两年间干了些什么。

在伍尔兹索普，23岁的牛顿在科学上做出了三大发现：微积分、光的色散和万有引力定律。在他逝世之前写的回忆中，关于引力定律的发现他是这样写的：

> 同年（1666），我开始考虑把地心引力延伸到月球轨道上……推导出使行星保持在它们的轨道上的力，必定与它们到回转中心的距离的平方成反比。由此，我比较了使月球保持在它的轨道上所需要的力与地面上的重力，并发现答案相当吻合，这一切是在1665和1666这两个疫症年代进行的，因为那些年代是我发现、思考数学和哲学的最佳年华。

首先，我们应注意到他说的“……那些年代是我发现、思考数学和哲学的最佳年华”。其次应注意到“答案相当吻合”这一关键词，它说明，他当时已经发现的月球在轨道上的加速过程，与根据平方反比定律推演出地面上物体的加速过程，即苹果下落过程，这两者是相当稳合的。牛顿似乎并不急于要进一步证明他预言的“答案”与实际“相当吻合”。的确，在发现自然界这样一个基本的定律过程中，他并未感到特别兴奋。事实上，后来有10年之久他完全没有考虑这一事件。

牛顿1667年初重返剑桥，1669年由于巴罗教授的推荐，接替了卢卡锡数学讲座教授的职务。

他回到剑桥后不久，牛顿完成了令他满意的光的色散的实验研究，

并设计制造了第一台消除色差的反射望远镜，因为当时的折射望远镜总是存在色差。但是，他把这些研究成果压了几年才发表。

牛顿根据新的原理制造出望远镜的消息不胫而走，人们迫切要求牛顿在皇家学会上展示这一望远镜。据说牛顿当时送去了两架望远镜，第二架在1671年的皇家学会上展出过。

1672年，牛顿被选为皇家学会的会员。或许是由于这一原因，牛顿答应了当时任皇家学会秘书的奥顿伯格（Oldenburg）的要求，在学会上介绍了他的发现，尤其是制造反射望远镜所依据的原理。在给奥顿伯格相继的两封回信中，牛顿写道：

> 我将以我卑薄的努力促进你们哲学计划的实现，并以此证明我竭诚的谢意。（1672年1月6日）

在第二封信中，牛顿建议报告他的光学发现，而不是对望远镜的描述，他写道：

> 让我讲解一个我不怀疑并且可以证实的哲学发现……而不是描述那架仪器，这将使我感到更加荣幸；在我看来，如果那不是迄今对自然的演变所做的最重要的发现，也是最有趣的发现。（1672年1月18日）

我应提醒读者注意“如果不是最重要的发现，也是最有趣的发现”这些话。这是牛顿第一次也是唯一的一次表达了他对自己的发现的热情。但是，当牛顿发表了关于光的色散实验的解释后，随之而来的却是一场灾难：一场激烈的论战爆发起来了。牛顿对那些批评者的无能，感到不可容忍的恼火，他们甚至对他已经用实验证明了的结论都不理解。缺乏理解是显而易见的，例如，惠更斯学派，甚至惠更斯本人都坚持认为：“用力学原理解释颜色组成的多样性，还存在相当大的困难，即使假定牛顿的关于白光分解为各色光的观点是正确的，这一困难仍然存在。”

起初，牛顿试图通过阐明他的方法去说服论战的对手：

> 最好和最安全的哲学研究方法似乎首先应该孜孜不倦地探究事物的性质，并通过实验确定事物的性质，然后通过相当慢的过程提

出假说去解释他们。假说只是帮助解释事物的性质而不是确定它们，除非可以用实验去证明。

附带说一句，我们也许注意到，这里牛顿已经道出了他后来正式提出的著名的格言：

我不杜撰假说。

牛顿未能从方法上说服他们。从此，他对科学出版、讨论、争论感到厌恶。他写信给奥顿伯格说：

我已经够了，因此决定今后只关心自己，而不再关心促进哲学计划的实现。（1672年12月5日）

我觉得我成了哲学的奴仆，一旦我从林纳斯（Linus）先生的事务中解脱出来，我将彻底地和哲学告别。除非为了我私自的满足，否则，它再也不会出现。因为我知道，一个人必须在两者之间作出抉择，要么决心什么新思想都不提出，要么成为一个捍卫新思想的奴隶。（1676年11月18日）

他这种对发表科学著作以及科学讨论、争论的厌倦感，在以后的许多年里又多次表现出来。有两段话是最好的例证：

能得到公众广泛的好评和承认，我并不认为这有什么值得我羡慕的。这也许会使与我相识的人增多，但我正努力设法减少相识的人。

拿起笔写那些可能会引起争论的文章是最可羞的事，这种想法与日俱增。（1682年9月12日）

光学发现发表后不久，牛顿引退了，以后10年他干了些什么我们知之不多。但我们知道，1679年牛顿证明了在中心平方反比引力的作用下，物体运动的轨道是一个椭圆，引力的中心在椭圆的一个焦点上。但是，他还是不公开这个结果。

多年以后，直到1684年，一次偶然的但非出自牛顿本意的事件导致了科学史历程的改变。1684年1月，雷恩（C.Wren）、胡克

（R.Hooke）和哈雷（E.Halley）在伦敦聚会时，他们提出了在平方反比引力作用下，行星的轨道是什么形状的问题。由于他们都无法解决这一问题，于是，哈雷于这年8月到剑桥去拜访牛顿，看牛顿对这一问题有什么看法。哈雷提出问题后，牛顿立即回答说：轨道是一个椭圆；而且他说早在7年前他就已经得出了这一结果。哈雷万分高兴，并希望看牛顿的证明。牛顿找了许久，但不知道把证明放在什么地方，牛顿答应，将重新证明并会很快寄给他。

对这一老问题的重新证明，似乎又提起了牛顿对整个领域的兴趣。到10月份，他已经解决了许多问题，足以把它们作为九个讲座的基础内容。这九个讲座是他在1684年下半年做的，讲座的题目是《论物体运动》，哈雷在收到牛顿寄来证明的同时，也听了牛顿的讲座。他又一次去剑桥，试图说服牛顿出版这些讲稿。

这时，牛顿的数学天才似乎被完全唤醒了，显示出娴熟的数学技巧。牛顿进入了最高度的数学活跃期。凭借自身的天才，凭借自己的意志和优势，牛顿坚强地向前推进，但这种推进是违反他的志愿和爱好的。最后，他终于完成了一生中智力上最伟大的成就，也是整个科学中智力上最伟大的成就。

先让我们暂停一会儿，来估价一下这一功绩的大小。牛顿出于自身的考虑，在1684年12月底，开始动笔写《原理》，17个月后，即于1686年5月把《原理》的三卷本的全部手稿交给了皇家学会。第一卷中有两个命题是他在1679年就已解决了的；第二卷中有8个命题是1685年6至7月解决的。第一卷总共有98个命题，第二卷有53个命题，第三卷有42个命题。因此，这些命题中的绝大部分命题都是在写三本书的连续17个月内宣布和证明的。除整个工作的规模宏大之外，完成的速度之快也是独一无二的。即使把《原理》中所完成的问题看成他一生事业和思想的结晶，牛顿在科学中的地位也仍然是无与伦比的。而且，用17个月的时间，就阐明、解决所有问题，并按逻辑的体系进行编排，这真是空前绝后的奇闻。人们能承认这一事件的唯一理由就是：它确实是这样发生的。

只有当我们知道牛顿取得了多么大的成就后，我们才会明白，拿牛顿与其他科学家相比是极不恰当的。事实上，只有莎士比亚、贝多芬才能与牛顿相提并论。

现在谈一谈《原理》的风格。与早期光学发现时表述思想的方法极不相同的是，《原理》是用冷漠的风格写成的，这种风格常使读者无周旋的余地。正如惠威尔（Whewell）贴切描述的那样：

> ……当我们读《原理》的时候，感到好像身在古代的军械库中，那里的武器尺寸如此之大，以至于当我们看到它们的时候，会不由自主地感到惊奇：能用它们作武器的是什么样的人？因为我们几乎提不动它……

显然，《原理》以刻板的、层叠的风格出现，很显然是经过深思熟虑的。因为在出版《原理》时，牛顿告诉德勒姆（Derham）说：

> 为了避免让那些在数学上知之甚少的人损害我的思想，我故意把《原理》写得深奥一些。但是，有才能的数学家，还是可以理解的。我想，他们理解了我的证明之后，会赞同我的理论。

尽管牛顿完成《原理》时年仅42岁，这时他的数学才能可以毫不夸张地说正处于高峰，并且在另一个40年里他完全可以保持这一才能，但是，以后他再也没有认真地进行科学研究了。他走向了另一条完全不同的生活道路。牛顿成了伦敦的重要人物，对所有的访问学者来说，他们肯定要拜会伊萨克·牛顿爵士。

牛顿是一个什么样的人呢？这一问题十分复杂，而且众说纷纭。尽管我们可以尽量概括，但有些个性特征是不能忽略的：对世事过于迟钝，对艺术缺乏兴趣，不能真正地理解别人。这些缺点大约是没有什么可争议的。

牛顿最杰出的才能也许是他专心致志的能力，正如凯因斯所写的：

> 他特有的才能就是，他能把一个纯粹的智力问题在头脑中持续保持下去，直到他完全搞清楚为止。我想他卓越的才能是由于他有最强的直觉能力和上帝赋予的最大的忍耐力……我相信，牛顿能把一个问题放在头脑中一连数小时、数天、数星期，直到问题向他投降，并说出它的秘密。

另外，正如狄摩甘（De Morgan）所说，他是

……如此沉浸在猜想的幸福中，似乎这样可以得到比任何证明方法得到的东西还多得多。

但是，牛顿生活中最主要的怪事就是他一贯故意地不显露他杰出的数学才能，而且对于他所做的那次超越任何人的贡献，也一贯不当回事。唯一的解释只能是，牛顿根本不认为科学和数学有多么重要。正如凯因斯所说：

……这似乎并不难于理解……这奇怪的精灵被撒旦引诱后相信，在三一学院期间，他解决了那么多问题，因而他可以凭借纯心灵的力量——哥白尼与浮士德结合体——解决上帝和自然界的一切秘密。最后，我不得不重复牛顿常常用来评价他自己的话：

我不知道我可以向世界呈现什么，但是对于我自己来说，我似乎只是像一个在海岸上玩耍的孩子，以时常找到一个比通常更光滑的卵石或更美丽的贝壳而自娱，而广大的真理海洋在我面前还仍然没有发现。

考虑到牛顿对别的事件不敏感和迟钝，有时人们不免要怀疑这一表述的诚实性。我不认为这样的怀疑是站得住脚的：仅有像牛顿这样的人，才能从他知识的高度，看到一个未被发现的“真理的海洋”。正如古代印度谚语所说：“只有大智大悟者才能探明智慧的源泉。”

Ⅳ

以上对莎士比亚、贝多芬和牛顿的创造模式的论述，虽然非常浅陋和不够充分，但有两个事实是十分明显的：一方面莎士比亚和贝多芬的创造模式惊人的相似，另一方面，他们和牛顿存在着明显的差异。这种相似和差异是不是偶然的呢？或者说，仅仅是由于这些大人物在人们心中留下了深深的印记，因而赋予了这种相似和差异现象的一般性呢？

把数学和诗放在一起来研究它们的创造性，也许能帮助我们说明问题。

本世纪杰出的英国数学家哈代在他的短文《一个数学家的自白》——它被斯诺（C.P.Snow）描绘成“对创造性思维最优美和空前绝后的写照”——中这样写道：

> 比起其它任何一门艺术和科学来说，所有教学家更不会忘记数学是青年人的游戏……伽罗瓦（Galois）21岁就夭折了，阿贝尔（Abel）27岁与世长辞！拉玛努扬（Ramanujan）33岁离开人间，黎曼（Riemann）只活了40岁。还有许多人如果不早去世的话，在以后的生涯中会做出更多更大的贡献……但是，我还从未看到一个数学家过了50岁还能取得重大的数学进展……一个数学家到60岁也许还有足够的竞争能力，但我们不可能期望他有独创性的思想。

对拉玛努扬的早逝，哈代还写道：

> 拉玛努扬的真正悲剧并不是他的早逝。当然，任何伟大人物的早逝，对人类都是一场不幸，但是，一个数学家到了30岁已经是比较老的了。因此，他的死也许并不像表面上那样是一场灾难……

除了哈代的这些话以外，我们再看看罗斯在马洛29岁逝世时说的话：

> 如果他还活着，有什么是他做不到的！——他的去世是英国文学最痛心的损失。

还有雪莱在30岁去世时，德斯蒙德—金赫勒（Desmond King Hele’g）说的话：

> 有一条规则，也许适用于任何诗人，那就是诗人最好的年华在30岁以后，这条规则对雪莱适用，同样对莎士比亚、米尔顿、华兹华斯、拜伦、坦尼森（Tennyson）也适用，事实上，英国的每一个活了30岁以上的伟大诗人都与这条规则相符。

然而，科学中的情况就大不一样了。托马斯·赫胥黎（T.Huxley）曾说：科学家过了60岁害多益少。

这些说法肯定会有争论，或者至少要求给予限定。但是，请考虑下面这一件事：

1817年，贝多芬47岁时，当一个较长的几乎没有写出什么作品的沉思时期快结束时，他对波特（Potter）说出了他的肺腑之言：“现在，我

知道如何创作了。”我不相信任何科学家过了40岁才说：“我现在知道如何研究了。”在我看来这正是不同的根源和核心。随着科学家的成长和成熟，他的无能也就越明显。

V

如果我们想找到科学工作者和艺术工作者之间的确存在的差异，并希望有一定程度的确定性，那么就必须对此做一定广度和深度的研究。但这已远远超出我的能力。但是，如果我不进一步举几个例子就此半途而废，似乎也不十分恰当。我再举科学史中的四个例子。

第一个例子是麦克斯韦，他被公认为19世纪最伟大的物理学家。麦克斯韦对物理学的主要贡献是建立了气体动力学理论和电磁场的动力学理论。麦克斯韦用电磁场方程（即麦克斯韦方程组）的形式引入的一些新的物理概念，每一个物理系的学生都了如指掌。这些新概念曾被爱因斯坦称为“是物理学自牛顿以来最深刻和最有成效的概念。”

在1860—1865年的5年间，这时麦克斯韦正值30～35岁。在伦敦国王学院任教授期间，他包含的两个领域贡献的四大部论文集就已经出版了。在这个紧张创作期结束时，麦克斯韦辞去了教授职位，引退回到苏格兰他的老家格伦莱尔（Glentair）（麦克斯韦的传记作者从来没有令人信服地“解释”过，为什么麦克斯韦认为他必须这样做）。在格伦莱尔麦克斯韦度过了6年的时光，在宁静的生活中，他似乎主要在着手写他的两卷本著作《电磁学通论》。1873年终于完成。1871年，经劝说麦克斯韦离开了格伦莱尔回到剑桥，又开始了学术生活，被任命为剑桥大学第一任卡文迪许实验物理学教授。1878年，麦克斯韦逝世，享年49岁。麦克斯韦在剑桥最后的8年中，主要致力于编辑出版卡文迪许的科学遗稿，组织和建设卡文迪许实验室以及处理其他各种各样的教学事务。麦克斯韦的早逝虽然是一个悲剧，然而我们也必须承认，他后来的工作没有达到他30岁前的高度。

第二个例子是斯托克斯（G.G.Stokes）。1849年，他刚过30岁就被选为剑桥的卢卡锡数学讲座教授。直到1903年去世，他一直担任这个曾由牛顿担任过的教席。斯托克斯是19世纪物理学和数学界最伟大的人物之一。他的名字一直与现今某些观点和概念联系在一起，比如，流体动力学中决定黏滞流动的纳维耶—斯托克斯方程；在黏滞媒质中决定小球

体收尾速率的斯托克斯定律（这一定律是确定密立根油滴实验的理论基础）；射电天文学上和大电流测量相关的特征极化辐射的斯托克斯参量；荧光波长一定大于激发光波长的斯托克斯荧光定律；另外，还有数学上的斯托克斯定理，这个定理除了是一个非常基础的定理之外，它还是当今微积分中微分形式发展的关键因素。

现在，斯托克斯的科学论文被收集在5卷本里。开始的3卷包括所有重要的基本概念和见解，是他在1842—1852年10年间的成果。后面两卷足以说明他后50年的全部科学成果。

哈奇森（G.E.Hutchinson，耶鲁大学的著名动物学家），他的父亲是斯托克斯晚年的好友，对斯托克斯作了出色的评价："斯托克斯（在卢卡锡教席上）非常可能在效仿他的伟大的前辈……牛顿做的，斯托克斯认为他也应该去做。"

我的第三个例子是爱因斯坦。1905年，无论对爱因斯坦还是对物理学界都是一个奇迹年。这一年爱因斯坦26岁，在这一年中他发表了三篇内容各不相同的划时代的论文。第一篇为狭义相对论奠定了基础，它的表述异常清晰、简洁和紧凑。第二篇合理地解释了分子布朗运动（与斯莫路霍夫斯基毫不相干）。第三篇将普朗克量子假设运用到逻辑极限处，形成了光量子的概念。紧接着10年，爱因斯坦一直迷恋于解决牛顿引力理论和狭义相对论的基本矛盾。牛顿引力理论假定力是一种瞬时的超距作用，而他的狭义相对论则以没有任何信号的传播速度能超过光速为前提条件。经过许多挫折和失败，爱因斯坦在1915年终于成功地建立了广义相对论。正如外尔（H.Weyl）后来所说的，爱因斯坦的广义相对论是"理论思维威力最伟大的范例之一。"

在广义相对论建立以后的若干年里，爱因斯坦对这一理论的许多分支作出了许多重要贡献，对统计物理的某些方面也作出了贡献。但是到了1925年，爱因斯坦对量子论的进一步发展不够关心。这一新的发展首先是由海森伯（W.Heisenberg）奠定的。在1927年的索尔维会议上，爱因斯坦的朋友埃伦菲斯特（P.Ehrenfest）对爱因斯坦说："爱因斯坦，我为你感到羞愧：你—直在反对新量子论就正如你的对手反对相对论一样。"海森伯悲哀地说，这位朋友的告诫只是一阵耳旁风。正如爱因斯坦的热切崇拜者兰佐斯（C.Lanczos）所说：

> 1925年以后，他对最新物理学动态的兴趣开始减弱。他自动放弃了他作为那时第一流物理学家的领袖地位，从原先研究的领域退出来，逃到他自愿去的流放地，这种状态只有少数几个同行愿意效仿。在他最后的30年中，他越来越像个隐士，和当时的物理学发展失去了联系。

我想举的最后一个例子是瑞利（Rayleigh）勋爵，他似乎不遵守哈代的一般规则。瑞利也许是经典数学和经典物理学中最伟大的支柱。在瑞利50年的科学生涯中，他的创造性自始至终具有惊人的稳定性和连贯性。他的科学工作体现在二卷本的《声学理论》和六大卷的《科学论文集》中。

1921年12月，J.J汤姆逊在威斯敏斯特大教堂所做的纪念演讲中，对瑞利的科学贡献做了如下的评价：

> 在构成这几卷著作的446篇论文中，没有一篇是无足轻重的，没有一篇不是把论述的课题向前推进的，没有一篇不是扫除了某种障碍的。在众多的文章中几乎找不到一篇因时代的进步而需要进行修正。瑞利勋爵以物理学作为自己的领地，拓展了物理学的每一个分支。读过他的文章的人都留下了深刻的印象，这不仅是由于他得到的新结果十分完美，而且在于它们十分清晰和明了，使人们对该主题有了新的领会。

这是一个令人注意的看法，谁有机会用到瑞利的《科学论文集》时，将会证明这一看法的精确性。

但是，为什么瑞利与爱因斯坦及麦克斯韦如此不同呢？也许从汤姆逊的同一篇演讲中可以找到答案。

> 有一些科学巨匠的魅力在于，他们对一个课题首先做出了说明，在于他们提出了以后被证明是很有成效的新思想。还有一些科学家的魅力则在于他们完善了某一课题，使该课题具有连贯性和明确性。我认为瑞利勋爵实质上属于第二类。

也许另外一个线索可以帮助我们了解瑞利。他的儿子（也是一位杰出的物理学家）曾问瑞利，对赫胥黎说的“一位过了60岁的科学家害多益少”这句话有什么看法。瑞利当时67岁，他的回答是：

如果这位年过60的科学家喜欢对青年人的成就指手划脚，那很可能害多益少。但是，如果你只做你所理解的事，那情况就可能不相同了。

我们也许能从中得到点启示。

Ⅵ

现在继续讨论一些具有共同性的东西。

因为艺术和科学都追求一个不可捉摸的东西——美，但艺术工作者和科学工作者具有不同的创造模式，这一点可以说使我苦思而不得其解。那么，美是什么呢？

在一篇极为动人的文章《精确科学中美的意义》中，海森伯给美下了一个定义。我认为这个定义是恰当的。海森伯的定义可追溯到古代，他说："美是各部分之间以及各部分与整体之间固有的和谐。"思考再三，我认为这个定义揭示了我们通常所说的"美"的本质。它同样适用于《李尔王》《庄严弥撒》和《原理》。

科学中有更多的证据表明，美常常是令人愉悦的源泉。在科学文献中，我们可以找到很多关于美的表述。这里我举几个例子。

开普勒：

> 数学是美的原型。

希尔伯特（D.Hilbert）（在纪念闵可夫斯基的致词中）说：

> 我们无比热爱的科学，已把我们团结在一起。在我们面前它像一座鲜花盛开的花园。在这个花园熟悉的小道上，你可以悠闲地观赏，尽情地享受，不需费多大力气，就能与彼此心领神会的伙伴同游尤其如此。但我们更喜欢寻找幽隐的小道，发现许多意想不到的令人愉悦的美景；当其中一条小道向我们显示这一美景时，我们会共同欣赏它，我们的欢乐也达到尽善尽美的境地。

外尔（Weyl）说（援引戴森的话）：

> 我的工作总是尽力把真和美统一起来，但当我必须在两者中挑选一个时，我通常选择美。

海森伯说（在一次与爱因斯坦讨论时）：

> 当大自然把我们引向一个前所未见的和异常美丽的数学形式时，我们将不得不相信它们是真的，它们揭示了大自然的奥秘。我这儿提到形式，是指由假说、公理等构成的统一体系。……你一定会同意：大自然突然将各种关系之间几乎令人敬畏的简单性和完备性展示在我们面前时，我们都会感到毫无准备。

以上这些说法也许显得过于笼统或者太一般化，下面我将用具体的、特殊的事例把它说得具体些、明确些。

毕达哥拉斯（Pythagoras）发现，在相同张力作用下振动的弦，当它们的长度成简单的整数比例时，击弦发出的声音听起来是和谐的。这是人们第一次确立了可理解的东西与美之间的内在联系。我想，我们会赞同海森伯这样一句话，毕达哥拉斯的发现是“人类历史上一个真正重大的发现。”

开普勒一定受到了毕达哥拉斯美的概念影响，当他把行星绕太阳的转动和一根振动弦进行对比时，他发现，不同行星的轨道有如天体音乐一般奏出了和谐的和声。开普勒深深感激上帝为他保留了这份发现，使他能够通过他的行星运动定律，得到了一种最高的美的联系。

较近的一个例子是伟大的科学家——海森伯的看法。在他发现了那把通向随后不断发展的量子理论大门的钥匙时，他记下了当一个伟大的真理显示出来的那个时刻对美的感受。

1925年5月底，海森伯患了花粉热，为了避开花草和田野，他来到了赫尔果兰岛休养。那里靠近海边，就在这一段时间里，他在解决当时量子理论的困难方面，取得了迅速的进展。他写道：

> 在短短几天内，我明白了在原子物理学中，只有用可观测量才能准确取代玻尔—索末菲的量子条件。很显然，我的这个附加假设已经在这个理论中引进了一个严格的限制。然后，我注意到，能量守恒原理还没有得到保证……

因此，我集中精力来证明能量守恒原理仍然适用。一天晚上，我就要确定能量表中的各项，也就是我们今天所说的能量矩阵，用的是现在人们可能会认为笨拙的计算方法。计算出来的第一项与能量守恒原理相当吻合，我很兴奋，而后我犯了很多计算错误。终于，当最后一个计算结果出现在我面前时，已是凌晨3点了。所有各项均能满足能量守恒原理，于是，我不再怀疑我所计算的那种量子力学了，因为它具有数学上的连贯性与一致性。刚开始，我很惊讶。我感到，透过原子现象的外表，我看到了异常美丽的内部结构；当想到大自然如此慷慨地将珍贵的数学结构展现在我眼前时，我几乎陶醉了。我太兴奋了，以致不能入睡。天刚蒙蒙亮，我就来到这个岛的南端，以前我一直向往着在这里爬上一块突出于大海之中的岩石。我现在没有任何困难就攀登上去了，并在此等待着太阳的升起。

在这一点上，请允许我发表一点个人感想。在跨越45个年头的整个科学生涯中，最令人震撼的就是，由新西兰数学家克尔（R.Kerr）发现的广义相对论中爱因斯坦方程的一个精确解，它为散布在宇宙中数量不明的大质量黑洞提供了极其精确的表示。这种“在美的面前震颤”，以追求数学美为动机的发现，竟然是大自然精致的复制品，这一难以置信的事实，使我不得不说美是人类思想最深层的反应。事实上，我想说的与此相关的每一件事，都可用拉丁箴言更简洁地表述出来：

Simplex sigillum veri ——简单是真的标志。

Pulchritudo splendor Veritatis ——美是真理的光辉。

Ⅶ

现在，我必须回到我的问题上来：艺术工作者和科学工作者的创造模式为什么存在差异呢？我不想直接回答这一问题，但我将作一些启发性的讨论。

首先，考察一下科学家和诗人相互间的看法如何。人们想到诗人对科学的态度时，几乎总会想到华兹华斯和济慈（Keats）以及他们那几句经常被引用的诗句。

一个摆弄手指的奴仆，

一心想窥探。

他母亲的坟冢？

一个自持有理夜郎自大的家伙，

一味凭借智力综合概括！

自然给我们带来知识的甜蜜；

我们的理智却胡折腾一气。

糟蹋自然的美丽外貌，

阴谋对她们解剖、分析。

（华兹华斯）

只要一触及冷漠的哲学，

一切迷人的东西都烟消云散？

天空绚丽灿烂的彩虹，

我们知道她为何模样那般。

安琪儿美丽的双翅，被哲学一触，

立即失去美的斑斓。

（济慈）

这些诗句也许与迪金森（L.Dickinson）的话遥相呼应：“科学出现的地方，文学就受到排斥。”

可以预料，科学家是不会赞同这些观点的。梅达沃（P.Medawar）反驳迪金森说：

> 我要找到的证据是，在文学降临之时，她驱走了科学……目前情形就是如此。简直没有任何可能使科学和文学相互补充，经过不懈努力达到一个共同的目标。相反，在期待它们相互合作的地方，它们却势不两立。

我反对这种互相指责的行为，因为这样只能使双方受到损害。所以，请允许我只说一句话：华兹华斯和济慈的态度并不具有代表性。科学家们对雪莱的态度倒是可以认真考虑一下。雪莱是科学家诗人，所以对雪莱的思想和工作最杰出的评论家是科学家金赫—勒，这决非偶然。金赫—勒指出："雪莱对科学的态度是强调他愿意生活在惊奇的现代思潮中"，雪莱"以英国诗歌史上无与伦比的准确性和精细性描绘了自然的作用机制。"怀特海（A.N.Whitehead）说：

> 雪莱对科学的态度和华兹华斯对科学的态度刚刚相反，他热爱科学，并在诗中一再流露出科学所提示的思想。科学思想就是他快乐、和平与光明的象征。

我将从雪莱的诗中选出两首诗，支持上述对雪莱的评论。第一首诗选自《云》[16]，这首诗"把创造的神话、科学专论和云彩快乐而传奇的历险故事有机地融合在一起"：

> 我是大地和水的女儿，
>
> 也是天空的养子，
>
> 我往来于海洋、陆地的一切孔隙——
>
> 我变，但是不死，
>
> 因为雨后洗净的天宇虽然一丝不挂，
>
> 而且，一尘不染，
>
> 风和阳光用它们那凸圆的光线，
>
> 把蓝天的穹庐修建，

我却默默地嘲笑我自己虚空的坟冢，

钻出雨水的洞穴，

像婴儿娩出母体，像鬼魂飞离墓地，

我腾空，再次把它拆毁。

第二首选自《解放了的普罗米修斯》，它被里德（H.Read）誉为“有史以来对人类追求智慧之光和精神解放的本性所作的最伟大的诗篇”。

闪电是奴仆，天穹至深处

抛撒群星，如绵羊一群，

数也数不清，群星在他眼前滚滚而过！

跨上暴风雨，驰过天穹；

地狱的呼喊，大白人间，

苍天，你还有秘密吗？人类已揭开面纱，一切都显露无遗。

现在我转到问题稍稍不同的另一方面。达尔文（C.Darwin）曾坦率地承认：

> 到了30岁，或更大一些，许多种类的诗，像弥尔顿的、格雷的、拜伦的、华兹华斯的、柯勒里奇的和雪莱的，都能给予我极大的快乐；我儿时曾沉醉于莎士比亚的戏剧中，尤其是他的历史剧……我还说过，以前绘画和音乐能带给我极大的愉快。但是。许多年以来，我没有耐心读完一行诗。后来，我试着读莎士比亚的书，感到单调乏味，味同嚼蜡，难以忍受。绘画和音乐也提不起我的兴趣……我的头脑似乎变成了一种机器，一种碾碎大量收积起来的事实，并使之变成一种普通规律的机器。但为什么会导致我大脑中较高级欣赏力赖以存在的那部分萎缩了吗？我实在不明白其中原因。

我们还可以考察法拉第电磁感应定律的发现。我们知道，这一发现形成了“力线”和“力场”的概念，这和当时流行的思想大相径庭。事实上，这些概念曾遭到当时许多人的冷眼。但是，麦克斯韦对法拉第的思想却独具慧眼，他曾预言道：

> 法拉第运用力线的思想来解释电磁感应现象，这一方法表明，他是一个具有很高水平的数学家——未来的数学家们可以从他那里得到有价值和富有成效的方法。我们甚至不知道怎么称呼我们正在努力建立和发展中的科学，这也许要出现另一位和法拉第一样伟大的哲学家才行。

然而，当法拉第讲述他对电的研究时，当时的财政大臣格拉斯通（Gladstone）却打断了法拉第的话，不耐烦地问道：“但它到底有什么用呢？”法拉第的回答是：“啊，阁下，也许不久你就会收它的税了。”法拉第的回答十分令人赞赏，因而常常被人们引用。

雪莱在《为诗辩护》中对科学耕耘所说的话，在我看来，对达尔文的坦白和法拉第的回答都是适宜的。雪莱说[17]：

> 科学已经扩大了人们统辖外在世界的王国的范围，但是，由于缺少诗的才能，这些科学的研究反而按比例地限制了内在世界的领域；而且人既然已经使用自然力做奴隶，但是人自身反而依然是一个奴隶。

你也许会认为雪莱对技术在近代社会中的作用麻木不仁，为避免这种错觉，我援引他接着说的话：

> 无庸置疑，从功利这种狭义的意义上说，提倡功利的人们在社会上也有他们应尽的义务。他们追随诗人的足迹，把诗人的种种创作中的素描抄写在日常生活的书本上。他们让出空间。他们给予时间。

雪莱的《为诗辩护》是英国文学史中最动人的文献之一。叶芝（W.B.Yeats）称它为“英语语言中对诗学基础最深刻的论述”。此文应全文通读，这里请允许我仅读几段。

> 诗，是最幸福最善良的心灵中最善良的瞬间的记录。

> 诗，可以使世间最善致美的一切永垂不朽；它捉住了那些飘入人生阴影中的一瞬，即逝的幻想……

> 真的，诗是神圣的东西。它既是知识的圆心又是它的周边；它包含一切科学，一切科学也必须溯源到它。它同时是一切其他思想体系的根和花朵。

> 诗人，是尚未被理解的灵感的祭司；是将未来的巨影投到现在的明镜，是表现了连自己也不解是什么的文字；是唱着战歌而又不感到何所激发之号角；是能动而不被动之力量。诗人，是未被世间公认的立法者。

在读雪莱的《为诗辩护》时，必然会提出这样的问题，为什么它和同样颇具天资的科学家写的《为科学辩护》毫无相似之处呢？也许当我提出这个问题时，我已对在报告中反复提出的问题做了部分回答。

演讲一开始，我就请求你们要有耐心，因为我讲的题目大大超过了我的理解能力。最后，请允许我援引莎士比亚《亨利四世》下篇的收场白来作我的收场：

> 第一，我的忧虑；第二，我的敬礼；最后，我的致词。我的忧虑是怕各位看了这出戏后会生气；我的敬礼是我应尽的礼貌；我的致词是要请各位原谅。

第4章　美与科学对美的探求（1979）

我们对于自然界的美都十分敏感，这种美的某些方面为自然与自然科学所分享，这不是没有道理的。但人们也许要问；在什么程度上追求美才是科学研究的目的？对于这个问题，庞加莱是一点也不含糊的。他在一篇文章中曾经写道：

> 科学家之所以研究自然，不是因为这样做很有用。他们研究自然是因为能从中得到了乐趣，而他们得到乐趣是因为它美。如果自然不美，它就不值得去探求，生命也不值得存在……我指的是本质上的美，它根源于自然各部分和谐的秩序，并且纯理智能够领悟它。

庞加莱还指出：

> 正因为简单和深远两者都是美的，所以我们特别刻意于寻求简单和深远的事实；我们醉心于探求恒星的巨大轨道，我们热衷于用显微镜寻觅极为细小的东西，我们欢欣于在遥远的地质年代中寻觅过去的痕迹，都是由于这些活动给我们带来了欢乐。

对于庞加莱的这些话，曾为牛顿和贝多芬撰写过杰出传记的沙利文（J.W.N.Sullivan）评论说：

> 由于科学理论的首要宗旨是发现自然中的和谐，所以我们能够一眼看出这些理论必定具有美学上的价值。一个科学理论成就的大小，事实上就在于它的美学价值。因为，给原本是混乱的东西带来多少和谐，是衡量一个科学理论成就的手段之一。

> 我们要想为科学理论和科学方法的正确与否进行辩护，必须从美学价值方面着手。没有规律的事实是索然无味的，没有理论的规律充其量只具有实用的意义，所以我们可以发现，科学家的动机从一开始就显示出是一种美学的冲动……科学在艺术上不足的程度，恰好是科学上不完善的程度。

著名的艺术批评家罗杰·弗赖（Roger Fry）在一篇名为“艺术与科学”的文章中，引用了沙利文的一段话后，颇有见解地指出：

> 沙利文大胆地说；“我们要想为科学理论和科学方法的正确与否进行辩护，必须从美学价值方面着手。”我想就这一点向沙利文提一个问题：一个无视事实的理论是否与一个符合事实的理论具有同样的价值？我想他将回答：否。然而依我个人之见，这个否定的回答并没有纯美学方面的理由。

关于罗杰·弗赖的问题，我在后面将提出一个不同的答案。但现在我要进一步讨论弗赖关于比较艺术家和科学家的冲动的观点。

> 从最单纯的感觉到最高的设计，艺术过程的每一步都必将伴随着欢快，没有欢快就没有艺术……同样，在思索中对必然性的认识通常也伴随有欢快的情绪，而且，对这种欢快欲望的追求，也的确是推动科学理论前进的动力。在科学中，不论是否有情感伴随它，关系的必然性依然同样地确定和可以阐明；而在艺术中，没有感情的激动，美学的和谐根本不会存在。没有激情，艺术中的和谐是不真实的……在艺术中，对关系的认识是直接的、有感情的——或许我们应该认为，它与数学天才的认识有惊奇的相似之处：数学天才们对数学关系具有直接的直觉，但要证明这些关系又超出了他们的能力。

现在我们从这些一般的论述转向具体的实例，看看科学家们如何体验美。

我的第一个例子与弗赖说的话有关，他说数学天才有时没有明显的理由就能感受到真理。1915年在数学上一鸣惊人的印度数学家拉玛努扬，想必你们有些人是知道的，他留下了大量笔记，其中有一本是几年前才发现的。在这些笔记中，拉玛努扬记下了几百个公式和恒等式，其中有许多最近才由拉玛努扬当时还不知道的方法证明出来的。华生（G.N.Watson）为证明这些恒等式耗费了几年时间，他写道：

> 研究拉玛努扬的工作和由此提出的问题，不禁使我

> 想起拉梅（Lame）在读埃尔米特（Hermite）关于模函数的文章时的评述：“令人惊心动魄。”而我无法用一句话说明我的感

受，像下面的公式

$$\int_0^\infty e^{-3\pi x^2}\frac{\sin h\pi x}{\sin h3\pi x}ax=\frac{1}{e^{2\pi/3}\sqrt{3}}\sum_{n=0}^{\infty}e^{-2n(n+1)\pi}$$

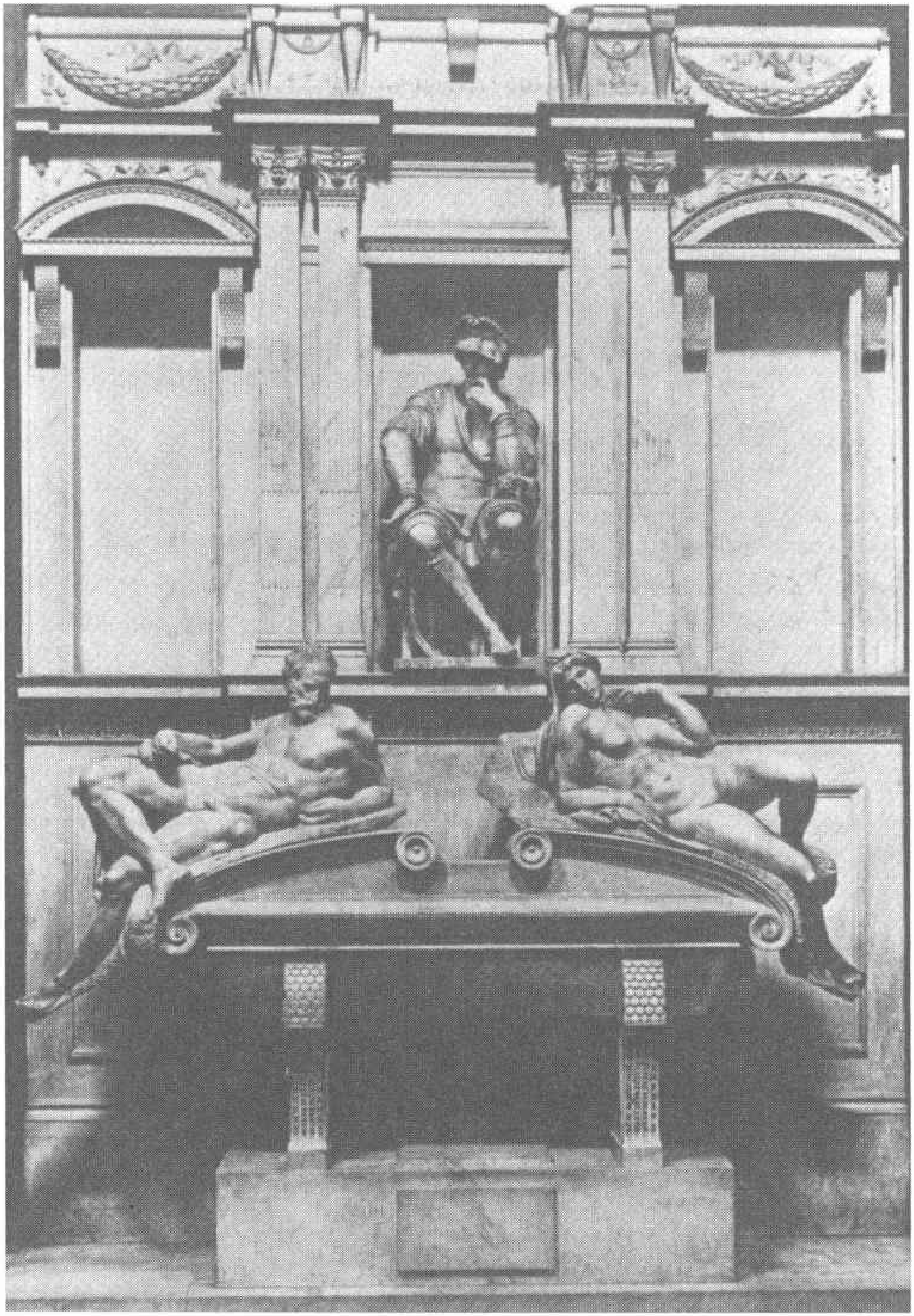
$$\times(1+e^{-\pi})^{-2}\times(1+e^{-3\pi})^{-2}\cdots(1+e^{-(2n+1)\pi})^{-2}$$

> 给我心灵的震颤，恰如我走进美第奇（Capelle Medicee）教堂的圣器室，见到米开朗琪罗的放在G.美第奇和L.美第奇墓上的名作“昼”、“夜”、“晨”和“暮”所引起的震颤一样。这两种感受是无法区分的。

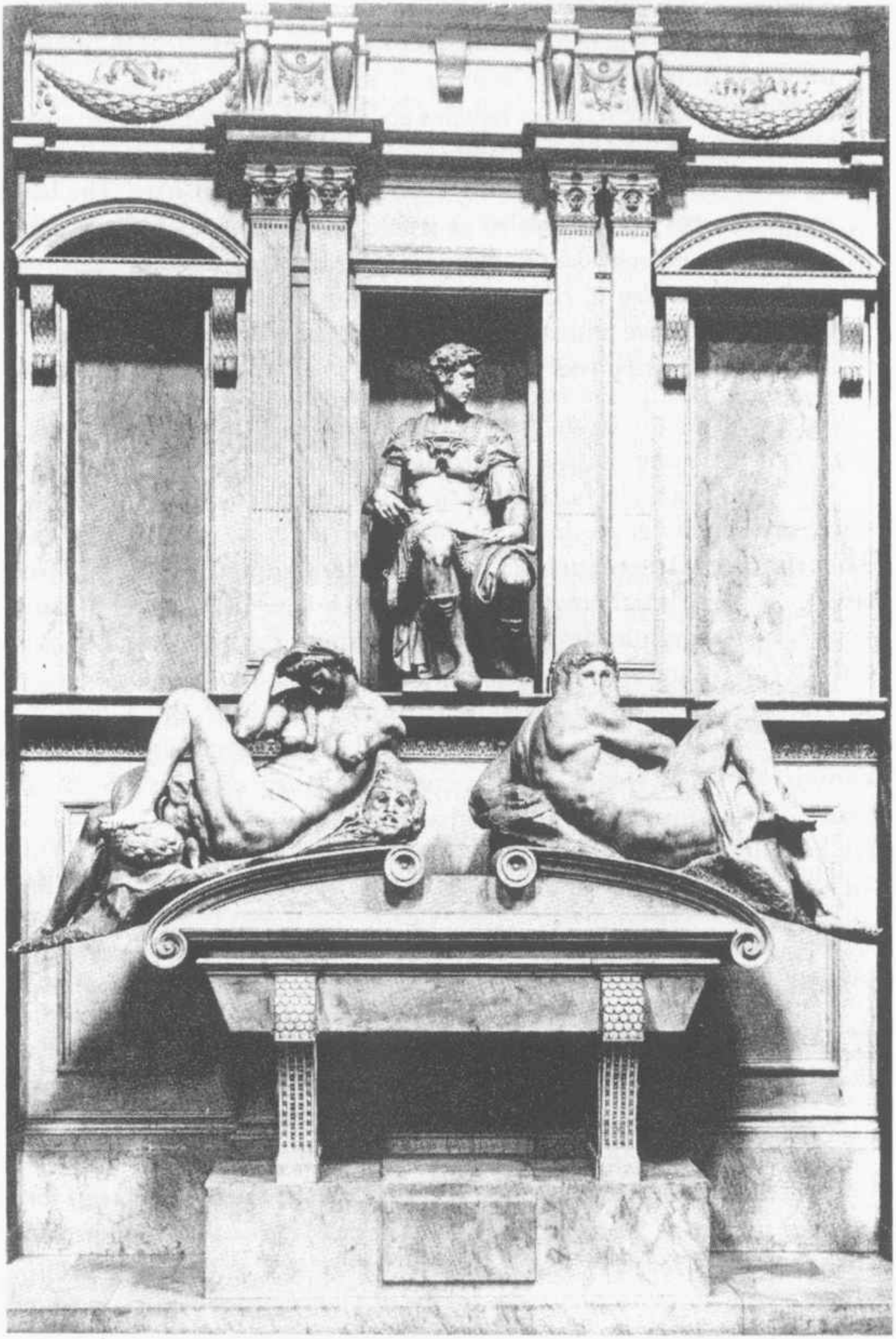
再举另一个很不相同的例子，这个例子说的是玻尔兹曼看了麦克斯韦论述气体动力学的一篇文章后的反响。在那篇文章中，麦克斯韦证明可以精确解出气体的输运系数，气体分子间的作用力是分子间的距离的负5次幂的函数。玻尔兹曼是这样说的：

> 一个音乐家能从头几个音节辨别出莫扎特、贝多芬和舒伯特的作品，同样，一个数学家也可以只读一篇文章的头几页，就能分辨出柯西、高斯、雅可比、亥姆霍兹和基尔霍夫的文章。法国数学家的风度优雅卓群，而英国人，特别是麦克斯韦，则以非凡的判断力让人们吃惊。譬如说，有谁不知道麦克斯韦关于气体动力学理论的论文呢？……速度的变量在一开始就被庄严宏伟地展现出来，然后从一边切入了状态方程，从另一边又切入了有心场的运动方程。公式的混乱程度有增不已。突然，定音鼓敲出了四个音节“令n=5”。不祥的精灵u（两个分子的相对速度）隐去了；同时，就像音乐中的情形一样，一直很突出的低音突然沉寂了，原先似乎不可被超越的东西，如今被魔杖一挥而被排除。……这时，你不必问为什么这样或为什么不那样。如果你不能理解这种天籁，就把文章放到一边去吧。麦克斯韦不写有注释的标题音乐。……一个个的结论

接踵而至，最后，意外的高潮突然降临，热平衡条件和输运系数的表达式出现，接着，大幕降落！



劳伦佐·美第奇雕像（正中壁龛内）与暮（左），晨（右）



朱理安诺·美第奇雕像（正中壁龛内）与昼（右），夜（左）

我举这两个例子是想强调，我们用不着刻意寻觅就可以发现科学中的美。但是，如果真要仔细寻求，我们可以得到最佳的例子。下面是其中的两个。

爱因斯坦的广义相对论，被外尔称之为推理思维威力的最佳典范，而朗道（Landau）和栗弗西兹（Lifschitz）认为，广义相对论大概是现有物理理论中最美的理论。爱因斯坦本人则在他的第一篇论述场论的论文结尾处写道："任何充分理解这个理论的人，都无法逃避它的魔力。"后面，我还将回头讨论这种魔力来自何处，现在我先将海森伯发现量子力学时的感受与爱因斯坦对自己理论的反应作一对比。我们有幸得到海森伯的自述，他写道：

> ……在只与可观测量打交道的原子物理学中，我逐渐明白了在原子物理学中，只有用可观测量才能准确取代玻尔—索末菲的量子条件。很显然，我的这个附加假设已经在这个理论中引进了一个严格限制。然后我注意到，能量守恒原理还没有得到保证。……因此，我集中精力来证明能量守恒定律仍然适用。一天晚上，我就要确定能量表（能量矩阵）中的各项……计算出来的第一项与能量守恒原理相当吻合，我很兴奋，而后我犯了很多计算错误。当最后一个计算结果出现在我面前时，已是凌晨3点了。所有各项均能满足能量守恒原理，于是，我不再怀疑我所计算的那种量子力学了，因为它具有数学上的连贯性与一致性。刚开始，我很惊讶。我感到，透过原子现象的外表，我看到了异常美丽的内部结构；当想到大自然如此慷慨地将珍贵的数学结构展示在我眼前时，我几乎陶醉了。

看了爱因斯坦和海森伯的这些有关自己发现的叙述，再回顾海森伯记下的他和爱因斯坦的一段对话，一定很有意思。海森伯记道：

> 当大自然把我们引向一个前所未见的和异常美丽的数学形式时，我们将不得不相信它们是真的，它们揭示了大自然的奥秘。我这儿提到形式，是指由假说、公理等构成的统一体系。……你一定会同意，大自然突然将各种关系之间几乎令人敬畏的简单性和完备性展示在我们面前时，我们都会感到毫无准备。

海森伯的这些话，与济慈的诗句遥相呼应：

美就是真，

真就是美——这就是

你所知道的，

和你应该知道的。

现在我想回到前面罗杰·弗赖提到的问题上，即如何看待一个在美学上令人满意但又认为它不真实的理论。

弗里曼·戴森（Freeman Dyson）曾经引用外尔的话："我的工作总是尽力把真和美统一起来；但当我必须在两者挑选一个时，我通常选择美。"我问戴森，外尔是否有具体例子说明他的这种选择？戴森说：有。引力规范理论是例子之一。这个理论是外尔在《空间、时间和物质》一书中提出来的。显然，外尔曾经承认这个理论作为一个引力理论是不真实的；但它显示出的美又使他不愿放弃它，于是为了美的缘故，外尔没有抛弃这个理论。多年之后，当规范不变性被应用于量子电动力学时，外尔的直觉被证明是完全正确的。

另一个例子外尔本人没有提到，但戴森注意到了。二分量中微子相对论性波动方程是外尔发现的，但由于它破坏了宇称守恒，物理学界有30多年没有重视它。结果，外尔的直觉再一次被证明是正确的。

因此，我们有根据说，一个具有极强美学敏感性的科学家，他所提出的理论即使开始不那么真，但最终可能是真的。正如济慈很久前所说的那样："想像力认为是美的东西必定是真的，不论它原先是否存在。"

确实，人类心灵最深处感到美的东西能在自然界得以成为现实，这是一个不可思议的事实。

凡是可以理解的也是美的。

我们也许会问：精密科学中的美在它被人们很好地了解和合理地阐明之前，怎么被认识到？阐明这种美的动力来自哪里？

这个问题自古以来就使许多思想家感到迷惑。正是在这一点上，海

森伯注意到柏拉图在《斐德罗》中表述的下述思想：

> 灵魂对美的光芒感到震惊，因为它感到灵魂深处有某些东西被唤醒了，这些被唤醒的东西并不是从外部输入的，而是一直潜藏在无意识领域的深处。

休谟（David Hume）在一句名言中表达了同样的思想：“事物的美存在于思考它们的心灵之中。”

开普勒发现了行星运动定律，他被这一发现所显示的和谐深深感动，在《世界的和谐》一书中，他写道：

> 人们可以追问，灵魂既不参加概念思维，又不可能预先知道和谐关系，它怎么有能力认识外部世界已有的那些关系？……对于这个问题我的看法是，所有纯粹的理念，或如我们所说的和谐的原型，是那些能够领悟它们的人本身固有的。它们并不是通过概念过程被接纳，相反，它们产生于一种先天性直觉。

最近，泡利（Pauli）更精确地表达了开普勒的这一思想：

> 从最初无序的经验材料通向理念的桥梁，是某种早就存在于灵魂中的原始意象（images）——开普勒的原型。这些原始的意象并不处于意识中，或者说，它们不与某种特定的、可以合理形式化的观念相联系。相反，它们存在于人类灵魂中无意识领域里，是一些具有强烈感情色彩的意象；它们不是被思考出来的，而是像图形一样被感知到的。发现新知识时所感到的欢欣，正是来自这种早就存在的意象与外部客体行为的协调一致。

泡利的结论是：

> 千万不要断言理性认识所建立的东西，是人类理性唯一可能的推测。

泡利所说的早就存在的意象与外部客体行为的协调一致，一旦被强烈地感受到，就会导致感受者对自己的判断及其价值坚信不移。否则，我们就无法理解一些伟大科学家们下述的言辞：

“热力学疯狂了。”热力学创建者之一开尔文勋爵（Lord Kelvin）在评论玻尔兹曼推导出维恩—斯忒藩定律时说。

“你从恒星的观点看；而我从大自然的观点看。”爱丁顿在和我的一次争论中说。

“正是在这一点上我不同意当今大多数物理学家的观点。”狄拉克在谈到量子电动力学中的重整化方法时说。

“确实，我们好像第一次有了一个巨大的框架，它足以包罗整个的基本粒子和它们之间的相互作用。我在1933年就有过的梦想由此得以实现。”1957年海森伯在谈到他与泡利合作研究统一场论时说（不过，这次合作的结果很不幸）。

“上帝不掷骰子。”这是爱因斯坦说的；他甚至还说过一句更带刺激性的话：“在评论一个物理理论时，我常问自己：如果我是上帝，我会不会这样来安排宇宙？”

爱因斯坦的后一句话，使人想起玻尔的劝告：“我们的责任不是规定上帝如何安排这个世界。”

也许我们正是应该从这些高度的自信中看出，有些伟大的人物也会有思想浅薄的表现。克劳德·贝尔纳（Claude Bernard）曾经说过：“过于自信的人不适于从事发现的工作。”显然，我踏入了一个危险地带，但这将使我有机会注意到一个曾使我极感迷惑的事实：它关系到两种非常不同的成长和成熟的方式，一种是伟大的作家、诗人、音乐家的方式，另一种是伟大的科学家的方式。至少在我看来，这两种方式有极大的差别。

当我们研究一个伟大作家或伟大作曲家的作品时，我们通常将它们分为早期、中期和晚期。而且，从早期、中期到晚期，这些作品总是经历一个日趋深刻和完美的过程。例如，莎士比亚和贝多芬，他们最后的作品是最伟大的。J.威尔逊（J.Dover Wilson）在叙述莎士比亚伟大的悲剧艺术时，曾非常精彩地描述了它的发展。

> 从1601年到1608年，莎士比亚沉浸在悲剧创作中；这8年中他走的路恰如一条山路，从平川开始，缓缓走上山坡，越往上路越

窄，到了顶峰，山脊如利刃，再往前则面临无底深渊。然后，立足不那么难了，再往下走，路又逐渐宽阔，最终落入另一侧的谷地。

八个剧本构成了这种悲剧的历程。首先是《裘力斯·恺撒》，它在悲剧期真正形成之前写成。这是一出并不邪恶但却软弱的悲剧。在《汉姆雷特》一剧中，邪恶势力出场了，它阴险、凶狠，但人性的软弱仍然占上风。在《奥塞罗》中，莎士比亚创造了第一个十足邪恶的人物形象伊阿古，同时，伊阿古的牺牲品是无罪的；莎士比亚不再让人性的软弱与上帝一同承担责任。《李尔王》把我们带到了万丈深渊之边，无穷尽的恐惧、无穷尽的遗恨，它终于铸就了世界文学史上最伟大的悲剧。

但莎士比亚并没有到此为止，他接着又写出了《麦克白》、《安东尼与克里奥佩特拉》（莎剧中最伟大的戏剧之一）和《科利奥兰纳斯》。威尔逊问道："在这些人类精神所承担的最艰险和最可怕的历险中，莎士比亚怎么解救自己的灵魂呢？"莎士比亚最终挣扎出来了，他的得救是由于他创作出伟大而绝妙的悲剧：《冬天的故事》《暴风雨》。

我不厌其烦地给你们讲莎士比亚的艺术发展历程，恐怕有些离题。但我之所以如此，的确是想向你们强调这一发展的重大意义。我相信，这种历程也适合于贝多芬的后期作品，其中包括《哈默克拉弗奏鸣曲》《庄严弥撒》，特别是最后的几首四重奏。

也许只有莎士比亚和贝多芬在他们的生命快结束时踏上了艺术的顶峰，并因此得救；也有一些人经历了相似的历程，他们由于坚持不懈的努力而逐渐攀上较高的山峰，只不过与莎士比亚和贝多芬相比较起来，不那么突出、显著。但对于科学家，我可就找不出有相似发展历程的例子。科学家最早的成就常常就是他们最后的成就。我这儿排除那些英年早逝的科学家，如科茨（coates）、伽罗瓦、阿贝尔、拉玛努扬、马约拉纳等，因为如果他们活得更长一些会有什么成就，我无法知道。科学家似乎都不能保持持久和连续的攀登，为什么会这样呢？我并不试图回答这个问题，我想转向一些更具体的思考。

现在我想讲的问题是，我们如何按文学或艺术批评的方式，把科学理论当作一件艺术品来评价。广义相对论为我们提供了一个好的例子，

因为几乎所有的人都同意它是一个很优美的理论。探求这种美的根源在哪里，我想是很有益处的。狄拉克有一句断言想取消这种探求，我认为是行不通的，他说：

> 〔数学美〕与艺术美一样是无法定义的，但研究数学的人要鉴赏数学美并不会感到困难。

我认为也不会有人满意玻恩（Born）的评语：

> 它〔广义相对论〕在我看来就像一件从远处观赏的伟大的艺术品。

（顺便直言两句，我不知如何理解玻恩的评语。难道广义相对论只能在远处欣赏吗？难道不能像其他物理学分支那样去研究和发展？）

尽管有许多固有的困难妨碍我们的讨论，但我还是试图阐明：广义相对论为什么会引起我们美的共鸣？我们为什么认为它是美的？为此，必须选定几条美的标准，我采纳了以下两条。

第一条是弗兰西斯·培根（Francis Bacon）的标准：

> 一切绝妙的美都显示出奇异的均衡关系。

（这里所说的“奇异”是指“感到非常意外，以致引起了惊讶和好奇”。）

第二条标准是对培根标准的补充，它是海森伯表述的：

> 美是各部分之间以及各部分与整体之间固有的和谐。

广义相对论显然具有奇异的均衡关系，因而符合培根的标准。这首先是因为广义相对论将一直被认为是完全独立的两个基本概念——时空和物质的运动，联系并结合起来了。正如泡利1919年所说：“时空几何不是既定的，它由物质及其运动决定。”随后，在引力与度规的融合中，爱因斯坦于1915年证实了黎曼在1854年所做出的伟大预言，即度规场与物质及其运动有必然的因果联系。

最大的奇异的均衡关系，也许就是时空关系的改变。正如爱丁顿所

说："空间不是许多聚在一起的点，而是许多相互连结的间隔。"

爱因斯坦广义相对论的建立，它的新奇也表现在另一个方面。即：我们可以欣然承认牛顿的引力理论需要修改，否则它将无法容纳光速的有限性和放弃瞬时的超距作用。承认了这一点，我们就可以推导出行星轨道与牛顿理论预言值的偏离是v/c，这儿v是行星在轨道上的速度，c是光速。在行星系中，这种偏离最大也不超过百万分之几。因此，如果爱因斯坦利用微扰法找到一个理论，允许牛顿的理论做出这样微小的修正，这就完全足够了。但这只不过是一种常规的方式，却不是爱因斯坦的方式。爱困斯坦要寻求一个精确的理论。他首先对物理性质作定性的讨论，然后将它与准确无误的数学优美性和简单性的感受相结合，就得出了场方程。爱因斯坦通过这种思辨性的推理思维，竟然得到一个完美的物理理论，这一事实很好地说明了外尔说的一句话，他说当我们跟随爱因斯坦的思想时，我们会感到"禁锢真理的墙已被推倒。"

上述议论只适用于导出场方程的理论基础。现在我们要进一步考察，这个理论是否符合美的第二条标准，即"各部分之间以及各部分与整体之间固有的和谐"。结果我们发现，这个理论的每一进程不仅显示出"奇异的均衡关系"，而且极其充分地满足了第二条美的标准。我对此要作点说明。

首先，广义相对论允许有黑洞的解。众所周知，黑洞把三维空间分为两个区域：内区和外区。内区是一个由光滑二维零表面（null-surface）为界的区域；外区是渐近平坦的，内区与外区不能相互沟通。有了这些非常简单和必要的限制后，出现了一个令人惊异的事实，广义相对论允许静止黑洞有一个单一的一族二参数解，这就是克尔族。克尔族的两个参数是黑洞的质量和角动量。更令人叹绝的是这个解族的度规是明确无误的，它轴对称，表示黑洞绕对称轴转动。

克尔几何的轴对称特性明显表明，作短程线运动的试验粒子其能量是守恒的，同时其绕对称轴的角动量分量也守恒。除了这两个守恒量以外，布兰登·卡特（Brandon Carter）还意外地发现，克尔几何允许试验粒子遵循第三个守恒定律。这样，支配实验粒子运动的哈密顿—雅可比方程其变量是可分离的；其短程线方程的解可以简化为求面积。这已经够令人惊讶了，但更令人惊讶的是所有的数理方程——标量波动方程、麦克斯韦方程组、狄拉克方程和支配引力波传播的一些方程，所有这些

方程在克尔几何中都可以分离变量，如像在闵可夫斯基几何中的情形一样，因而可以得到明确的解。

当我们领悟到彭罗斯（Penrose）和霍金（Hawking）的奇异性定理的要求时，我们必将感到同样的震惊。彭罗斯和霍金的奇异性定理要求我们的宇宙必然起源于一个奇点，这样，我们不得不思考一些令人难以置信的物理过程，在这种过程中密度、体积线度和时间间隔的数量级分别为10^{93} g/cm^3 、10^{-33} cm和10^{-44} s！

霍金定理表明：黑洞的表面积总是在增长，这暗示黑洞的表面积与热力学的熵具有同一性；这就导致热力学、几何学和引力之间有着密切的联系。

所有这些都显示出奇异的均衡关系。

以上我说的一切，都符合我作为出发点的两条美学标准。但还有一个方面我们应该考虑。

当亨利·莫尔（Helary Moore）在10年前访问芝加哥大学时，我曾问他应该怎样看雕塑：是站远一点还是靠近一点看比较好。莫尔回答说，最伟大的雕塑能在任何距离上进行观赏，因为在不同的距离会显示出不同的美。莫尔还以米开朗琪罗的雕塑作品为例作了说明。同样，广义相对论在我们探讨它的每个层次上，都显示了奇异的均衡关系。举一个例子就足够了。

如果把爱因斯坦方程扩展到爱因斯坦—麦克斯韦方程，即适合于空间充满电磁场的方程，并寻求一个球对称的解，我们就可以得到一个描述有质量和电荷的黑洞的解。这个解是赖斯纳（Reissaer）和诺德斯特姆（Nordstr m）发现的，是众所周知的史瓦西（Schwarzschildl）解的一个推广。由于黑洞有电荷，所以一个电磁波如果射到黑洞上，有一部分显然会以引力波的形式反射回来。反过来，如果引力波入射到黑洞上，一部分引力能量也会以电磁波的形式反射回来。令人惊讶的是，在任何频率的情形下两种反射回来的部分都一样。这个结果是出乎意料之外的，其根本原因目前被归结为（经典）物理学定律的时间可逆性。这个例子表明，广义相对论在探索的每一个层次上都显示出奇异的均衡关系。正是这一事实，使广义相对论具有无与伦比的美。

至此，我的评论都只限于已被接受的伟大的思想，它们出自伟大的心灵。但我们不能由此认为，只有伟大的心灵在伟大的思想中才能感受到美。同样，创造的欢乐也不仅仅只限于少数几个幸运的人。事实上，只要努力去领会均衡的奇异性和各部分之间以及各部分与整体之间的固有的和谐，我们都有机会体验美和创造的欢欣。除此以外，把一个科学领域的研究对象和谐地组织起来，使它有序、规范、连贯，我们同样也可以得到满足。这样的例子很多，例如雅可比的《动力学讲义》、玻尔兹曼的《气体理论讲义》、索末菲的《原子结构和光谱》、狄拉克的《量子力学原理》和薛定谔晚年撰写的一些珍贵的解说性论著。正如古希腊罗马哲学家普罗提诺（Plotinus）所说，透过物质现象隐约可见的永恒的光辉，在这些书中像彩虹那样展示在我们眼前。

最后，我认为我们每个人可以用我们自己的方式，在追求科学美中得到满足，正如弗吉妮娅·伍尔芙的《浪》中的演员们一样：

> 这儿有一个正方体，那儿有一个长方体。演员们将正方体放到长方体上，放得非常精确，想做一个完美的住处。从外面还看不出什么名堂，但结构可以看出来了；虽然明摆着不完美。我们虽不多才多艺，但也不那样无能；我们已经作出了长方体，并把它们放到正方体上。这就是我们的胜利，这就是我们的安慰。

第5章　米尔恩讲座　爱德华·阿瑟·米尔恩和他在现代天体物理学发展中的地位（1979）

I

米尔恩是在1921年进入天体物理学领域的。那时，天体物理学才刚刚起步，对现代天体物理学的两大支柱——恒星大气理论和恒星结构理论还少有研究。

在1920年，天体物理理论方面的书籍很少，可以说仅有一本，即罗伯特·埃姆登（Robert　Emden）的《气体圈》（Gaskugeln）。这本书出版于1907年，对引力场中处于平衡态的气体的质量作了极为详细的研究；并且指出，平衡态下气体的压强与气体密度的某种乘方成正比。这就是多方气体圈理论，它在爱丁顿和米尔恩的研究中扮演很重要的角色。埃姆登的著作除了包括广为大家所知的多方气体圈理论之外，还讨论了太阳大气中的物理条件。这一讨论实际上是对卡尔·史瓦西的如下推论作了解释：太阳外层不可能处于对流平衡，而是处于辐射平衡。史瓦西是从太阳边缘的有限亮度分析中提出上述推论的，下面我马上会谈到这一点。同一时期，天体物理学的另一重要成就出现在阿瑟·舒斯特（Arthur　Schuster）的论文中。在该论文中，舒斯特处理了与太阳和恒星大气吸收谱线有关的热辐射传递理论问题。

史瓦西于1914年对辐射平衡的概念做了更深入的分析。1916年，爱丁顿将这些概念引入到作为一个整体的恒星系的热平衡中去，并由此着手关于恒星内部结构的首篇论文。1918年，爱丁顿的恒星亮度变化的脉动理论问世。

当时，原子论仍处于幼年时期，萨哈（Saha）关于在恒星大气中估计的温度和压强下，各种原子激发和电离的论文即将发表。

米尔恩就是在这个时候进入天体物理学领域的。我得马上插一句，他的贡献主要还不在于与他联系在一起的科学的具体进步，而更多的在于他的治学态度和治学作风。后面我会详细地讲一讲。

Ⅱ

很幸运的是，米尔恩将注意力转移到天体物理学领域时，他首先遇到的问题正好与他的治学风格和处理问题的方法合拍。正如我将要指出的，米尔恩刚开始进入天体物理学领域的研究成果，多年来一直保持领先地位，并且为我们了解恒星外层的一些不变的特性提供了证据。正是由于这一点，我想细致地谈一下他在这方面的工作。

米尔恩开始研究的问题是如何解释太阳不同区域亮度不同的现象，也就是如图1所示的太阳边缘发暗的现象。不仅从太阳不同区域发出光线的总亮度不一样，而且从太阳不同区域发出的不同波长的光（或说不同颜色的光）的亮度也不一样。

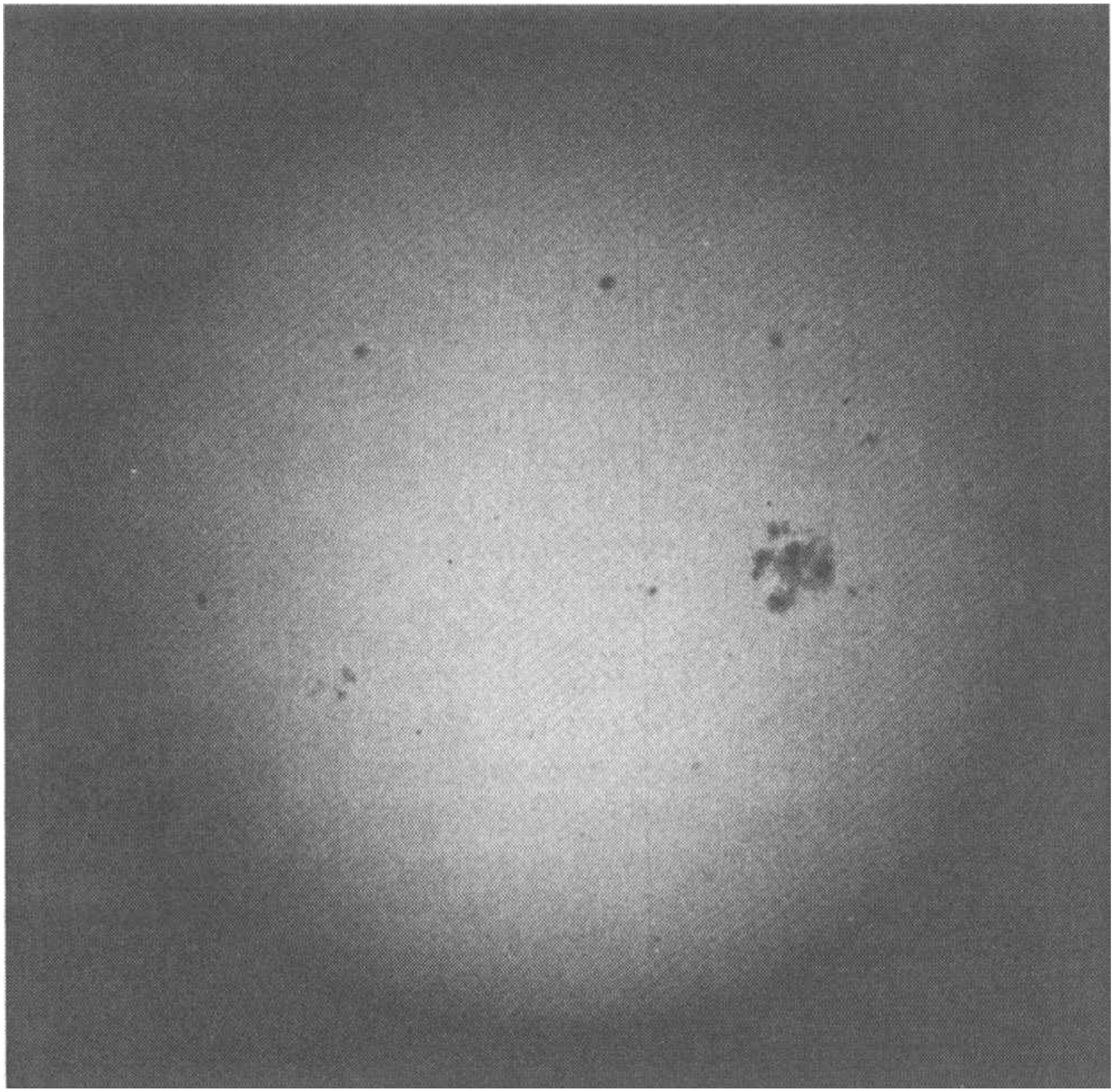


图1太阳边缘发暗现象照片

图2很清楚地表明：太阳边缘发暗现象简明地反映了辐射强度与辐射角的关系。1906年，史瓦西研究了太阳边缘发暗问题，将它与辐射平衡优势和由此产生的太阳外层大气温度变化联系起来。这样，问题的解释就变得很简单了。

一个基本的事实是物体的辐射来自物体的不同深处，但由于外层物质的阻挡，越是深处其辐射衰减的程度就越厉害。基于这个事实，我们可以说物体的辐射反映了物体表面下某一平均深度的辐射特征。这个平

均深度对于各种波长和各个角度的辐射都是一样的，我们可以通过考察覆盖层对辐射的衰减程度来测量它。换句话说，我们在任何情况下总是能有效地观察到一个单位的光学深度（辐射穿过单位光学深度的物质其衰减因子为1/4）。

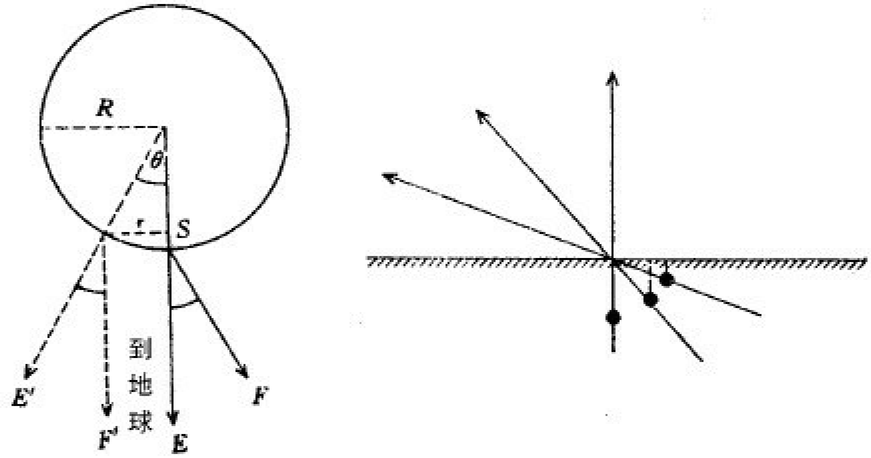


图2辐射角和大气中的温度梯度引起发暗现象示意图

有一些辐射是以一定角度斜着穿过大气的，显然它们所代表的辐射深度一定小于垂直于表面的辐射所代表的辐射深度。既然我们认为越深层温度越高，那么以一定角度发出的辐射，其温度要比垂直辐射的温度低。所以，倾斜辐射的强度要比垂直辐射弱（见图2）。或者说，恒星的边缘必然发暗。

通过以上说明，可以清楚地看出要解释边缘发暗现象，必须解决的基本问题是恒星外层的温度分布是怎样的。一旦温度的分布弄清楚了，就可以直接将任何给定角度的辐射强度同物质对各种波长光线的不透明度（即吸收系数或者吸收能力）联系起来。

设$\bar{k}$ 代表平均吸收系数，τ代表用$\bar{k}$ 量度的光深，Tτ为深度τ处的温度。根据普朗克分布可知，在深度为τ处辐射的频谱分布为：

$$B_v(T_\tau) = \frac{2hv^3}{c^2} \frac{1}{\exp(hv/kT_\tau - 1)} \qquad (1)$$

这里，v、c、h分别为频率、光速和普朗克常数。相应地，与法线

方向成θ角，频率为v的辐射的强度为：

$$I_v(\theta)=\int_0^\infty d\tau B_v(T_\tau)(\frac{k_v}{k})\exp\left[(-k_v\sqrt{k})\tau\sec\theta\right]\sec\theta \qquad (2)$$

其中，k_v 是给定频率下的吸收系数。

显然，通过比较实际观测到的辐射强度与方程（2）得出的辐射强度，可以推得吸收系数随波长（由 $k_v\sqrt{k}$ 计算）而变化；这种变化明显规定了某种与太阳大气的构成有重要关系的东西。

以上我提到的问题，米尔恩在他早期的论文中作了彻底的解决。他从对太阳的观测数据中，推导出太阳连续吸收系数随波长变化，如图3所示。

米尔恩强调，他导出的变化关系有两个特征：第一，在整个可见光谱部分，吸收系数逐渐增大，在约8000Å附近达到一个极确定的最大值；第二，超过8000Å后。连续吸收系数开始减少，在16000Å附近存在一个最小值。

在随后的几十年中，许多人用另外的方式重复了米尔恩的分析，均证实了米尔恩的一些主要结论。下面我将谈其中的一个证实方式，它是由1946年（米尔恩后约25年）查隆格（chalonge）和考格诺夫（Kourganoff）的研究所提供。

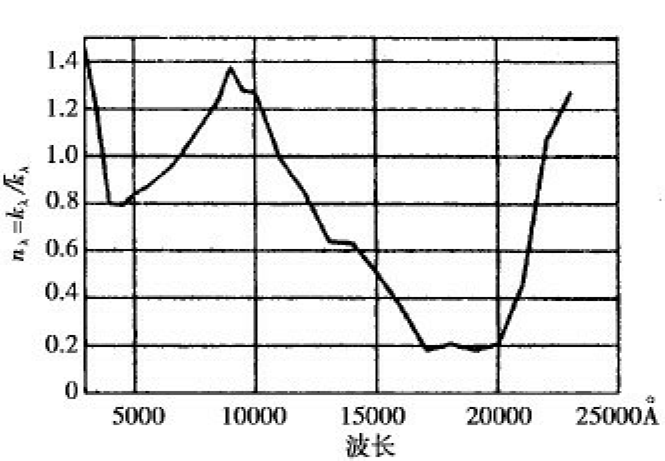


图3米尔恩推导出的太阳连续吸收系数随波长变化图

假定某一深度处温度为T，试求覆盖层对各种波长辐射的不透明度。显然，米尔恩的基本理论可以解决这个问题。查隆格和考格诺夫对该问题的分析结果如图4所示，它和米尔恩的结论符合得很好。

导出连续吸收系数随波长变化这一论断，其推理分析的简洁性表明，它涉及了太阳大气的基本成分。至于成分可能是什么，只有在20世纪40年代人们能够明确分离出太阳大气的一种基本成分以后，才能得到答案。

这段历史代表着米尔恩最早最基础的研究成果，请让我追溯一下这段历史。

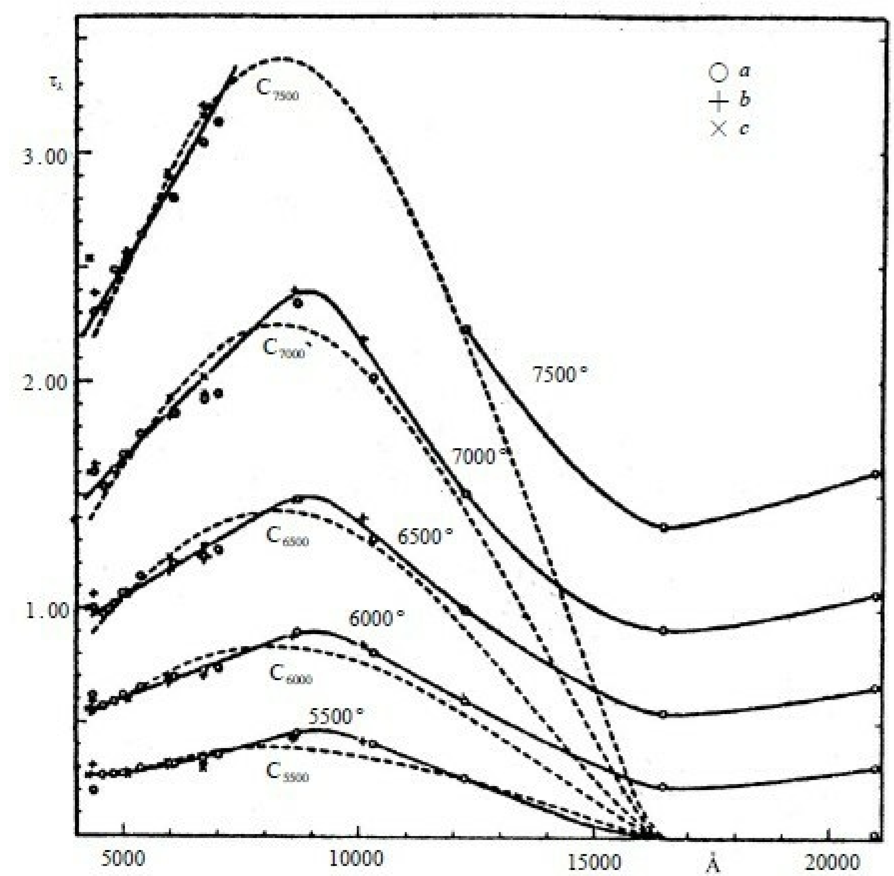


图4不同波长对应的光球层的光深（τ_λ）〔查隆格和考格诺夫（1946）〕，虚线代表H^-吸收；右边部分代表自由—自由跃迁

利用变分法可以确定原子系统基态能量的上限。1930年，海勒拉斯

和贝特各自独立地利用变分法揭示出：氢原子可以结合一个电子形成负离子，其结合能稍大于3/4电子伏特。但是，只有到了1938年，威尔特（Rupert Wildt）才回头研究这个16年前由米尔恩提出来，而后一直被回避的基本问题。威尔特指出，如果氢确实像其它证据表明的那样丰富的话，那么负氢离子一定是非常集中地存在于太阳大气层中。这是一个富有成效的建议，它是使后来恒星大气理论有可能取得进展的关键。但是，在能够得出负氢离子是导致太阳大气连续吸收谱的原因之前，还有几个困难必须克服。

主要困难是从理论上判定负氢离子的吸收系数是否是连续变化的。在这儿现在不适于谈这个问题解决的历史。那些对此感兴趣的人，可以参考最近贝茨（David Bates）在《物理报道》（Physics Reports）上发表的一篇详细论述文章。图5（摘图中可以清楚地看出，负氢离子连续吸收系数再现了1922年米尔恩曲线中的几个基本特征（参看图4）。不可忽视的是，那时候负氢离子从理论预言是一个原子种类，但对这一理论预言的实验证明还需等待10年。

米尔恩的研究是现代天体物理学伟大的开端。关于这个问题讲到这里为止。

III

下面转入天体物理学发展的另一阶段。大约在米尔恩致力于太阳连续光谱研究的同时，福勒（R.H.Fowler）和达尔文（C.G.Darwin）致力研究的统计力学新方法正取得进展。那时候，萨哈在《哲学杂志》和《皇家学会会刊》上发表了一系列论文，首次将统计平衡理论（更确切地说，是热力学）定量地应用于恒星反变层（stellar reversing layers）。

萨哈的理论建立在下面的观察基础上：恒星大气的每一稳定电离态吸收不同的光谱，事实上是吸收一组不同的谱线。因而，在任何恒星光谱中，连续光谱的吸收谱线的相对强度给出了不同电离态下的反变层中原子相对数量的某种信息，同时也就给出了温度和压强这两个状态参量的有关信息。萨哈早期应用这种思想，是基于对特征谱线的第一条和最后一条这样两条边缘谱线位置的分析。他指出，在这两个位置处，反变层中能吸收该特征谱线的原子的浓度一定非常小；如果能估算出相应的压强值，就可以推算出温度值。

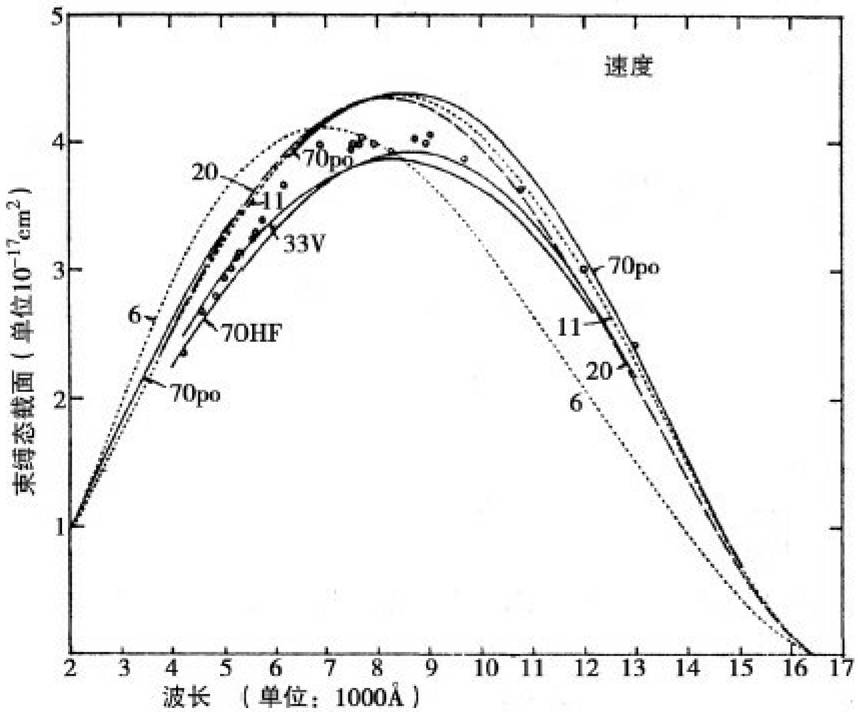


图5H$^-$ 的光致分离截面。实验数据取自史密斯（Smith）和伯奇（Burch）1959年的研究成果。束缚态波函数的变分参数值标在每条曲线上。虚线表示放射出的电子用平面波近似的情况：实线表示放射出的电子作更精确近似的情况：70po（po代表极化轨道），70HF（HF代表哈特里—福克展开），33v（v代表变分，由简化的科恩—费斯巴赫方法和罗顿堡—斯坦因束缚态函数得出）。所有的计算均依据速度矩阵元

这些早期计算的精度值得怀疑，因为对特征谱线的边缘谱线出现的条件做精确描述是很困难的，我们不知道与边缘谱线对应的“非常小的”原子浓度究竟有多小；而且边缘谱线出现的位置还依赖于产生这种谱线的元素的相对丰度。福勒和米尔恩在1923—1925年间发表的一些论文中，对此问题重新阐述如下：

> 在其他条件都相同的情况下，恒星光谱中某给定吸收谱线的强度，总是随反变层中能吸收该谱线的原子浓度的变化作同样变化。

这样，就避开了有关谱线边缘出现的问题。福勒和米尔恩首先致力于确定恒星序列中某条谱线达到最大强度时的位置。在所述前提下，最大强度发生在能吸收谱线的原子的浓度最大处；为此所需的条件仅仅只有温度和压强。也就是说，一定压强下，谱线强度最大时的温度可以从平衡态性质简单地推导出来。这是第一种利用萨哈的理论定量处理恒星温度和压强问题的令人满意的方法。利用这种方法，福勒和米尔恩首次为哈佛巨系光谱型建立了一种理论温标，这是天体物理学的一个真正的

里程碑。

随后，米尔恩又写了一些论文论述平均压强和平均温度的概念，这两个概念是他和福勒研究的基础，反过来又需要至少在两个方面重新定义。首先，必须明确“吸收谱线的强度”这一短语的确切含义；其次，要考虑到温度、压强的变化以及产生吸收谱线的吸收层的各种物理参数的变化。事实上，我们必须构造出恒星大气模型。尽管米尔恩指出在重新定义平均压强和平均温度的过程中，一些基本因素应全部考虑到，但他却没有对此做任何进一步的分析。更进一步分析和完善的工作落在潘尼柯克（Pannekoek）、安索尔德（Unsold）、明纳尔特（Minnaert）及其他一些人身上。建立恒星大气模型现在已成为一个科研热点，这一切均起源于萨哈、福勒和米尔恩在基本物理概念方面做出的不朽的努力。

Ⅳ

接下来我要谈的是米尔恩在恒星结构方面的研究工作。早在致力于恒星大气理论研究的同时，米尔恩就已经关注恒星结构理论方面的问题了。在1923年发表的一篇论文中，他考察了低速转动对爱丁顿标准恒星模型和质光关系的影响。这是一篇非常优秀的论文，它将数学和物理学的方法紧密结合起来〔请让我插一句，正是米尔恩的这篇论文促使我10年后创立了完善的畸变多方理论（theory of distorted polyrtopes）〕。

但是，直到1929年，米尔恩才真正将注意力转移到恒星结构问题的研究上来。由于同爱丁顿激烈的争论造成的压力的影响，开始并不怎么顺利。据我看，这场争论是一个令人不愉快的插曲，它对米尔恩以后的研究工作产生了非常不利的影响。在这里我不想对此争论评头论足，但是，无论怎样解释1929年以后米尔恩的工作，都不可能不涉及这段历史。

1926年，福勒在一篇奠基性的论文中指出，白矮星（如天狼星的伴星）内部物质的状态不可能是一种准确服从方程P=RρT的理想气体（式中P代表压强，ρ代表密度，T代表温度，R代表普适气体常数）；它应该服从由当时新提出的费米—狄拉克统计给出的状态方程，而且是它的极限形式，即低于某一临界值的自由电子的能级均被占满，高于此临界值的能级均未被占据。换句话说，物质处于“简并”状态。

福勒的讨论充分证实，爱丁顿的假设——恒星完全是气态的，服从常规状态方程——不具普适性，上面提到的白矮星就是一个反例。白矮星中物质是简并的，压强和浓度的关系几乎与温度无关。因而，探索在什么时期和在什么条件下恒星内部物质发生简并，这在当时是一个顺理成章的研究课题。但是米尔恩的研究不是这样直接展开的。他是从如下的前提（至少他认为这是个先验的结论）开始的：所有恒星都存在简并区域，而且所有恒星可分为两大类：中心凝聚结构型和坍塌结构型。这两类恒星的区别主要在于简并区域的大小。

1931年，米尔恩发表了详细讨论这个问题的第一篇论文。在论文中，米尔恩提出了一些很有效的分析方法，用于建立复合恒星（Composite Stellar）结构。他认为在这种复合恒星结构的不同区域中，压强和浓度之间的关系是不同的。此外，米尔恩还鼓励他的老友福勒系统地研究支配多方分布的埃姆登微分方程的全部解。

〔插一句话，我想引用哈代在1931年1月举行的一次皇家天文学会的会议上的讲话，也许你们有些人记得，哈代20年代在牛津大学任Savillian教席的几何学教授。他（如像他后来向我承认的那样，当时心无诚意地）说：

> 作为一位数学家，关于恒星到底是什么的两种观点的争论，我一点也不关心……但是我对福勒先生的论文有特殊的兴趣……他的论文可能是论文集中唯一有永久价值的，因为他肯定是对的，因而其他人的观点很可能被证明是错误的。

恐怕哈代的预言大部分已经被证实了吧！

我们还是回到米尔恩的研究上来。米尔恩向皇家学会递交第一篇论文之前，他就已经被告知，完全简并的恒星其质量不能超过某个极限值；同时该事实又为恒星简并核所能包容的质量设置了一个上限；最后，考虑到大质量恒星中辐射压越来越重要，巨大质量的恒星肯定不可能产生简并区。但是，米尔恩没有接受这些事实，相反，他写道：

> 如果量子力学的结论与许多显而易见的推理发生矛盾的话，那么，状态方程所依据的基本原理和先前提到的普遍原理，其中必然有一个是错误的。开尔文关于太阳引力年龄的计算看来是很完美的，但是它与许多当时还未认识到的判断相矛盾。我很清楚，物质

的行为不可能像你们描述的那样……你们罗列出一群声名显赫的权威，如玻尔、泡利、福勒、威尔逊等，这的确给人以深刻的印象，但这并不能引起我的兴趣。

今天看来，米尔恩的消极态度显然使他不能认识到，正确利用费米简并可以直接导出如下的事实：质量巨大的恒星在耗尽能量之后，必须坍缩为黑洞。这个结论是由爱丁顿提出的，但又为爱丁顿本人和米尔恩所不愿接受。他们的失败说明了靠自己的信念去考察世界是十分危险的。

正如我前面说过的，米尔恩在处理复合恒星模型的过程中，提出了一些很有效的分析方法。他的方法从理论上讲很适合处理具有简并核的恒星模型，恒星在它们各自的质量上限内，肯定会有简并核。米尔恩本应该很容易地完成这方面的工作。但是他没有做，这对于米尔恩本人和天体物理学来说都是不幸的。

Ⅴ

对运动相对性和宇宙学的研究，是米尔恩最后阶段也是工作量最大的工作。在谈到这些研究之前，我想简单提一下1935年米尔恩在一篇论文中对恒星运动学的精彩分析。在论文中，米尔恩仿照斯托克斯的分析方法，分析了恒星系内部可能存在的差动运动（differential motions）。斯托克斯曾将流体动力学中流体运动分为三种运动形式：转动、剪切和扩张。从这种分析观点出发，恒星基速度的双正弦波（随银河径度而变化）具有一个同恒星间距离成正比的振幅，就显然不证自明了。米尔恩的分析为以后更多的动力学讨论奠定了基础。

Ⅵ

下面我要谈米尔恩科研工作的最后一个方面，他自己确认这一方面是他对科学做出的最重要的贡献。1943年7月6日，他在给我的信中涉及了他的宇宙膨胀理论：

我不知道我是否曾向你谈到过我对这种理论的看法，只知道它的理论结构与一般的数学物理有着惊人的不同。但我确信，一旦我的理论被人们认识后，它将被视为一种革命。这样吹捧自己的工作

实属少见，但我内心的确是这么认为的。

也许将米尔恩先生仅对我个人讲的话透露出来是不合适的，但就我的看法而论，我不能对他的理论一味地去赞扬，对此我必须坦率地承认这一点。

在阐述他的运动相对论时，米尔恩坚决主张引力理论没有广义相对论也能够很完善。他为一本论文集写了一篇题为“没有相对性的引力”的文章，这本论文集在1949年被送给爱因斯坦。后来米尔恩的论文又被收集在《在世哲学家文库》第七卷《爱因斯坦：哲学家和科学家》（P.A.希尔普编）中。作为答复，爱因斯坦在该书末篇中写道：

> 对米尔恩先生坦率的见解，我只能说他们的理论基础太狭隘。依我来看，一个人如果不利用广义相对论，他就不可能在宇宙学领域里通过理论研究获得任何可靠的成果。

与爱因斯坦的这一观点相对照，我们再列出米尔恩对广义相对论的看法：

> 用黎曼度规描述现象给出了建立概念的基础，爱因斯坦的引力理论无论如何也不能算是从这一概念基础导出的必然结论。我从未相信过它存在的必然性……广义相对论如同一个开满鲜花同时又杂草丛生的花园，近来期望的鲜花盖过无用的杂草而使花园美丽壮观。

米尔恩接着说，“在我们的花园里只有鲜花开放”。

为了完整起见，我想谈谈自己的观点。广义相对论建立在下面的假设基础之上：一种引力理论应用于“小范围”的物体上时必须能还原为牛顿定律，比如处理太阳系中物体的运动；而且一个理论只有当它同其他物理定律相一致时（如纳入等效原理），它才能被推广到宇宙更大的范围中去。而米尔恩建构理论的顺序刚好与此相反。他假设我们可以先构造出宇宙理论，用它解决引力问题，然后再下推到小范围的现象中。米尔恩没有实现他的设想，也许这种设想本来就是不可能实现的。

好啦，现在你们已经听到了三种不同的权威观点。

虽然我从总体上否定了米尔恩的这段工作，但我应当立即指出，他的工作中有一些关键处仍闪耀着创造性的光辉。比如，米尔恩利用观察者交换光信号来分析洛伦兹变换就是准确而又经济的典范。应该让更多的人了解它，正如邦迪（Bondi）指出的：

> 我觉得，我们都深受米尔恩的教益，但对这一点说得不够。他在宇宙学的研究中，引入了雷达测距法的设想。

接下来我想谈谈米尔恩的宇宙学思想，这些思想在现代科学中占有稳固的位置。

在20年代末和30年代初，作为统一宇宙理论基础的事实有2个：

1.在一级近似下，河外星云的分布是局部均匀和各向同性的。

2.星系在不断地远离我们，不同星系之间也在不断地远离，远离的速度与它们之间的距离成正比，就像哈勃定律指出的那样。

弗里德曼（Friedman）和勒迈特（Lemaitre）建立的相对论模型受到了欢迎，特别是受到了爱丁顿的欢迎。但此理论对上面两个事实的讨论有意无意地给人一种印象：只有广义相对论才能将上述两个事实融为一体。这就夸大了广义相对论的作用。米尔恩正确地指出，对上述2个事实有一个很简单的解释，并不一定需要引入任何特殊的理论。

观察到的宇宙膨胀和哈勃关系仅仅表明，我们现在所能观察到的星云曾经在某一时刻聚集在一个狭小的空间内。假设在时刻t，星云聚集在一个狭小空间内（如图6），星云的速度方向各异但大小相等，设为V。经过一段足够长的时间间隔（$t-t_0$）之后，这些彼此相同的星云将向外分散运动，并局限在一个比较薄的球壳内，球壳的厚度为V（$t-t_0$）。假设原来的狭小空间内除了包含速度大小为V的星云外还含有速度大小为V/2的星云，那么经过时间间隔（$t-t_0$）后，这些星云将局限在一个厚度为V/2（$t-t_0$）的球壳内。更一般地说，如果原始狭小空间内包含有各种速度大小的星云，经过一段足够长的时间后，速度大小各异的星云将会分散开来，也就是说，它们将分处在离中心距离不等的地方。并且遵守哈勃定律。用米尔恩的话说："物以类聚，人以群分。"

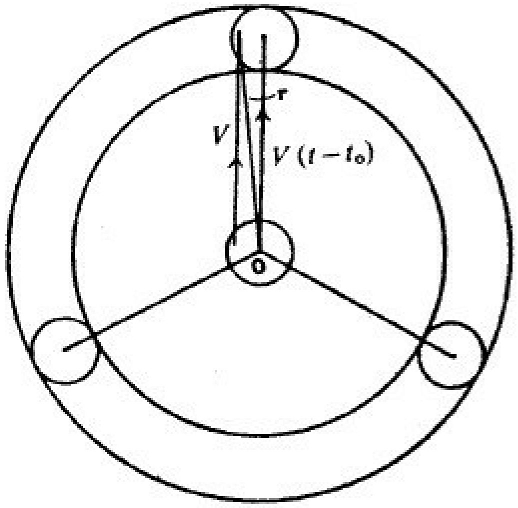


图6对于一个开始局限于一个狭小空间内的系统，哈勃关系产生示意图

我们上面描述的简单模型表明，观察到的宇宙膨胀不过是宇宙初始高平均密度的一个结果罢了——今天没有人反对这一结论。

另外，正如米尔恩强调指出的，一个同质的粒子系统，所有的粒子都相对于某一取定的粒子作后退运动，其速度与它们到该粒子的距离成正比。这样的系统具有一个非常值得注意的特征：如图7所示，根据速度矢量合成的平行四边行法则可以证明，只要不太靠近边界，取系统中的任何粒子作为参考粒子，对系统中运动的描述效果都一样。

在这样的系统中，每个粒子都可以视自己为中心，其他粒子都在径向上离它而去，其速度与它们到中心粒子的距离成正比，并具有相同的比例常数。换句话说，宇宙就是一个同质和各向同性，内部的运动满足哈勃关系的系统，因而宇宙这样一个大系统就应该和上述粒子系统具有类似的特征：宇宙各星系具有共同的起源。即以任何一个星系作参考系来描述宇宙，效果都是一样的。米尔恩把这后一结论作为宇宙学原理（cosmological principle）提出来。他认为该原理是至高无上的，不容违背的，是他的运动相对性理论的核心。

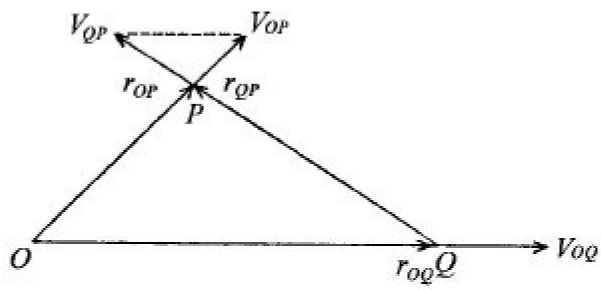


图7图示粒子的速度与粒子间距离成正比的系统满足宇宙学原理。粒子P和Q均从O点开始运动，速度与它们距Q点距离成正比。但是，对于一个随粒子一起运动的观察者来说，粒子P唯一的可见运动是远离Q运动，速度与P、Q间的距离成正比

由于前面我已提到过的原因，我不想对米尔恩的运动相对性理论作任何进一步的评价。但是，我应该讲清楚，根据米尔恩的想法，如何从宇宙学原理和牛顿定律推导出一个局部恰当的，而且同弗里德曼相对论宇宙模型相符合的宇宙描述。

显然，宇宙学原理要求对每一个观察者来说，宇宙相对于他自己都是球对称的。根据一个在牛顿理论框架和广义相对论理论框架中都有效的原理，如果一个系统中物质的分布相对于原点具有球对称性，那么处在球边界上的粒子将仅受球内部物质的吸引作用。这样，只要扩张的速度远小于光速，压力对惯性的影响可以忽略，我们就可以利用牛顿引力定律和牛顿的概念体系分析该系统的动力学性质，而且可以期望这样处理的结果在所述的限制范围内对广义相对论的理论框架来说，也是有效的。的确，如米尔恩，更确切地说如米尔恩和麦克雷（McCrea）指出的，从牛顿理论分析得出的方程和从相对论分析得出的方程是一致的，当然也是在上述限制范围内。若超出了这个范围，即压强对惯性的影响不能不考虑，膨胀的速度可以和光速相比拟，那我们就必须借助于广义相对论了。

我不知道米尔恩是否同意我对他的思想的表达，是否答应我向广义相对论作出的让步。然而，我上面主要根据米尔恩的思想描述的理论，今天已是每个宇宙学的学生都必须学习的内容。

Ⅶ

从总体上看，米尔恩是怎样的一位科学家呢？

首先，最重要的一点是他善于将复杂的问题分解成一些基本要素，然后对每个要素逐个地分析其内涵和意义。干脆利落、精力充沛的著述风格是他敏锐分析能力的标志。他曾告诉我，他的笔常常跟不上一泻千里的思路。此外，他力求将分析性的问题解决得优美精致，并以此为乐。他的这一令人羡慕的特点洋溢在他所有著作中，尽管在后期的作品中这一特点被多少掩盖了一些，但那是出于自卫和争论的缘故。一旦气氛比较自由，思维不受约束了，他就会在他那奔流的思维中，以及他的论证的过程和构架中，把显而易见的愉悦传递给读者。尤其是在他妙不可言的著作《矢量力学》一书中，这种愉悦的感受可说达到了顶点。读者还可以从其他一些论文中得到相似的感受，虽然这些作品可能不是他科研工作的主流，但这些珍品以许多方式反映了米尔恩最优秀的一面。

如果让我选择一篇文章能体现米尔恩的创造力、他的风格和他对自己所做事情的酷爱的话，我将选择他于1933年发表在《牛津季刊》上的论文“非稳态动力学及其在造父变星光变上的应用”（就是连这篇文章也受到了一些不必要争论的影响）。1933年我们在我的剑桥寓所里谈话时，米尔恩向我讲述了包含在这篇文章中的思想。他想知道，造父变星的光变现象怎样才能用一个普通的理论框架而不需要参考任何内部特殊参量（比如温度和压强等）来解释。他说，造父变星说到底就是一部热机。根据他从纽沃尔（H.F.Newall）那里学过的关于格里菲思热机的知识，他很快建立了一个理论框架，由光变相对振幅的时间导数得出下面函数方程：

$$k\varphi(t)+\varphi[t+b+\varphi(t)]=0 \qquad (3)$$

其中，k和b是两个常数。

方程（3）有许多值得注意的性质。例如，如果在某时刻φ为零，那么在一个时间间隔为b内的无限多连续瞬间，φ均取零；如果k取1，则方程有周期解，周期为zb；此外还有其他一些有趣的性质。

后来，米尔恩在一次皇家天体物理学会的会议上谈到这一工作时，曾不无自豪地说，既然方程（3）

给出的周期解再现了造父变星光变曲线的一些特点……那么建

立一种分析（造父变星光变）理论，从中导出方程（3），不是不能做到的。

要对米尔恩作全面的评价，必须考虑到他在剑桥的早期岁月被第一次世界大战中断；1923年得了重病，这病于1950年夺去了他的生命；他一生中还掺杂着重大的个人悲剧：科研工作蒙受两次世界大战的影响，停滞多年；晚年又病魔缠身；更重要的是，和爱丁顿的争论使他的科学经历倍添苦涩。如果我们考虑到这些不利因素，再牢记他做出的许多实实在在的成绩，那么我们就会像他的老朋友和同事哈里·普拉斯克特（Harry Plasktett）那样公正地说："他虽死犹生，他的智慧永存。"

第6章　纪念A.S.爱丁顿诞辰一百周年讲座（1982）

（1）爱丁顿：当代最杰出的天体物理学家

I

1944年11月，爱丁顿在他62岁的时候离开了人世。在大西洋彼岸，与他同时代的伟大人物罗素（H.N.Russell）为此写道："爱丁顿爵士逝世，天体物理学因此失去了自己最杰出的代表。"[1]　我这次讲座的头一讲，就是从罗素那里得到了启发。

在评价爱丁顿对天文学和天体物理学的贡献之前，我打算先作一些传略性介绍，这样，对爱丁顿个人风格可以有一个初步的印象。

阿瑟·斯坦利·爱丁顿（Arthur Stanley Eddington）1882年12月20日出生在威斯特摩兰的肯特尔。父亲阿瑟·亨利·爱丁顿（Arthur Henry Eddington）是肯特尔镇斯特拉蒙加特学校的校长及校董。道尔顿（J.Dalton）100年前曾在那里任教，48年后爱丁顿被授予肯特尔镇荣誉镇民，他当时回忆说：

> 我父母婚后在肯特尔生活了一段不长的时间，但肯特尔的传统已深深印在我童年的记忆中。肯特尔一直将科学工作认作为一项最重要的公共服务，这不是从任何物质意义上，而是基于它对社会作出的贡献。我为此深感欣慰。肯特尔更早的时候就与科学和一位伟大的化学家联系在一起，这位或许是历史上最伟大的化学家，曾经是斯特拉蒙加特学校的校长，一百年后我父亲成为这同一所学校的校长，我也出生在那里。从道尔顿开始我们有了原子，我自己现在也成为一名原子研究者。道尔顿一定留下一些思想的胚芽，它们在斯特拉蒙加特延续。我喜欢想起那种连续性，能够沿着肯特尔伟大的科学家开辟的道路前进，我为此深感自豪。[2]

爱丁顿的父亲死于1884年，他的母亲带着两个年幼的孩子，斯坦利

和比他年长四岁的姐姐温尼弗雷德，迁往滨流韦斯顿。在这里，爱丁顿很早就显示出对大数的迷恋；他学会了24×24乘法表，有一次开始数《圣经》有多少字。爱丁顿从未失去对大数的兴趣。在以后的生活里，写到天体大小和距离时他常喜欢明确地带上所有的零。比如，1926年爱丁顿在牛津给英国学术协会做晚会演讲时是这样开头的：

> 恒星具有相当稳定的质量，太阳的质量为——我把它写在黑板上：
>
> 2000 000 000 000 000 000 000 000 000吨。
>
> 但愿没写错数字零的个数，我知道你们不会介意多或者少一两个零。可大自然在乎。[3]

1935年爱丁顿的兴趣已转向天体宇宙，他是这样介绍这门学科的：[4]

	英里
太阳的距离	93 000 000
太阳系范围（冥王星轨道）	3600 000 000
最近恒星的距离	25 000 000 000 000
最近星系的距离	6 000 000 000 000 000 000
宇宙的原始圆周长	40 000 000 000 000 000 000 000

谈到大数，最著名的当然要数他1939年出版的《物理科学的哲学》第11章开头一句：

> 我相信宇宙中有15 747 724 136 275 002 577 605 653 961 181 555 468 044 717 914 527 116 709 366 231 425 076 185

631 031 296个质子和相同数目的电子。[5]

这个数为136×2^{256}。后来人们称它为爱丁顿数。罗索问过爱丁顿，是他自己算出这个值还是让别人算出来的。爱丁顿说在一次跨越大西洋的旅途中自己动手算出的。

1893至1898年，爱丁顿在滨流韦斯顿的布里米林学校上学。记得爱丁顿曾告诉我，他在学校玩的一种游戏就是像刘易斯·卡罗尔那样编造出语法正确却无实际意义的英文句子。有一次他对我说起一个例子，“站在篱笆旁，听起来像一只萝卜（To stand by the hedge and sound like a turnip.）”。后来在他更严肃一些的著作中，爱丁顿习惯于引用这样的句子。他的斯沃思莫尔演讲“科学与未知世界”中就有这样的句子：“人类个性无法用符号来估量，正如你无法摘录一首十四行诗的平方根一样。”[6]

我不想更深入地谈及他的童年教育，只想提到一个事实：他1898年进入曼彻斯特的欧文学院，一呆就是四年。他在欧文学院的老师包括舒斯特爵士和拉姆爵士（Sir Horace Lamb）。爱丁顿似乎对拉姆一直怀有深深的敬意。20年代初，爱丁顿已是英国科学界最著名的人物之一，据说他曾经说过：“尽管我明白被人称作狮子般勇猛意味着什么，但我宁愿是一只羔羊（Lamb）。”

在曼彻斯特度过这段重要岁月后，爱丁顿于1903年依靠一份初级入学奖学金进入剑桥继续攻读，这份奖学金后改为中级奖学金。1904年他成为数学学位考试一等及格者；1907年获得史密斯奖学金，同年入选三一学院。1936年在哈佛大学的一个午餐会上，我有幸坐在怀特海（Alfred North Whitehead）旁边。怀特海是1907年的投票人之一。他回忆说他当时选择了爱丁顿，尽管另一位提交的论文篇幅长得多。怀特海回忆起此事时显得很自豪。

1907年，也就是被选入三一学院的同一年，爱丁顿应皇家天文学家克里斯蒂爵士（William Christie）之邀，进入格林尼治天文台任高级助手。在这个位置上干了5年，直到1912年被选为剑桥的普卢米安讲座教授，接替乔治·达尔文爵士。罗伯特·伯尔爵士（Robert Ball）1914年去世后，爱丁顿还担任了剑桥天文台台长。从此他在这个显赫的职位上一干就是30年。在剑桥，他先是与母亲和姐姐同住，后来又单独和他姐姐

住在一起。

Ⅱ

结束传记性介绍前，我想谈一谈爱丁顿的一般观点和习惯。爱丁顿是教友派信徒，作为一个教友派信徒，第一次世界大战期间，他是一个基于道义的反战者。下一次讲座里还要详细介绍他对那场战争基于道义的反对态度。这里我只想说说他1929年的斯沃思莫尔演讲“科学与未知世界”。爱丁顿在这篇演讲里极其诚恳地表述了他对宗教、科学和生命的看法：

> 宗教信条是一种巨大的障碍，它阻碍着科学家的观点与宗教所经常要求的观念达成一致……探索精神激励着我们，它拒绝将任何形式的信条作为探索的目标。如果发现有那么一所大学，那里的学生们需要完成的一项练习就是吟诵自己对牛顿运动定律、麦克斯韦方程和光的电磁理论的笃信和忠诚，我相信你一定震惊不已。即便背诵的是我们自己钟爱的理论，或者是近几年新的理论，我们依然为此遗憾。如果教育学生应将这些结果当作需要背诵和认可的东西，学生们就根本理解不了科学训练的意图。科学可能达不到自己的理想；而且尽管冲突的危险不至于以这种极端的形式表现出来，可真要在信条和教义面前坚持我们的立场，通常并不容易做到，对通俗科学来说也是如此。
>
> 拒绝信条与保持生活信念并不矛盾。科学没有信条，并不等于我们对信仰漠不关心。我们的信念不在于相信已有关于宇宙的全部知识永远正确，而是确信我们仍在前进途中。如果说我们所谓的事实是变化着的阴影，那么这阴影正是永恒真理之光投下的。
>
> 信念与盲目自信是完全不同的两回事。[7]

有一件小事。爱丁顿的朋友们都知道他非常喜欢在春秋两季独自骑车旅行。可也许只有少数人知道他对旅行作了细心记录。我1936年12月离开剑桥前，爱丁顿曾给我看过一大张英国巴塞洛缪旅行地图，上面布满他仔细描出的前些年旅行经过的不同路线。他还告诉我摊开看的那张已是第二张地图，他养的狗把前面那张撕烂了，害得他重弄一张地图，并将第一张地图上的全部路线抄录下来。

爱丁顿还告诉过我，在格林尼治任高级助手时，他和西德尼·查普曼（另一位旅行酷爱者）曾订出一条衡量旅行记录的指标。该指标指的是一个最大天数N，在不同的N天里骑车行走N英里以上的路程（我后来有次向查普曼提起这个指标时，他已记不起来；但他记得自己经常和爱丁顿比较旅行记录）。

爱丁顿后来给我的每封信都要谈及最新的N值，这也许令人感动。以下摘自他的两封信：

> 我旅行的N值还是75。今年复活节运气不佳，骑车旅行两次，可只有74.75英里，没有用。还需另外四次骑车旅行才能完成下次跃迁。然而，那几天天气和地方都不错，特别是南威尔士……明天我得穿上神秘的衣服——短裤和丝质长筒袜！还有英王授予的勋章。（1938年7月4日）

> 现在N值为77。我记得你在时还是75。最近的跃迁发生在几天前，那天我在沼泽地里骑了80英里。由于无法得到夜宿的地方，1940年以来一直没有骑车旅行，所以记录前进得很慢。（1943年9月2日）

最后一件事。爱丁顿特别喜欢做《泰晤士报》和《新政治家和国家报》上的纵横字谜。解开一个字谜他很少有超过5分钟的。爱丁顿有时让我看着他解字谜，速度之快令人惊讶。

III

允许我现在转而评价爱丁顿对天文学和天体物理学的贡献。1906年爱丁顿开始投身天文学时，格罗宁根的卡普坦（J.C.Kapteyn）已作出一项革命性的发现。卡普坦是研究恒星运动的一位伟大先驱。他的发现如下：

在那以前，人们一直以为，恒星处于一个局部静止系统中，其运动完全是随机的，不存在任何偏离的方向，这就是说在这种系统之中，其周围恒星的平均速度为零。关于恒星本征运动和径向速度，它们所揭示的一个基本问题就是确定太阳的运动，也就是确定太阳在其周围恒星所决定的静止系统中的“奇异速度”。根据速度不存在偏离方向的随机性假

设，本征运动分布投影到小天区时应为一个拉长的椭圆[见图1（a）]。但卡普坦看到的并非如此，他发现是一条双叶曲线[见图1（b）]。

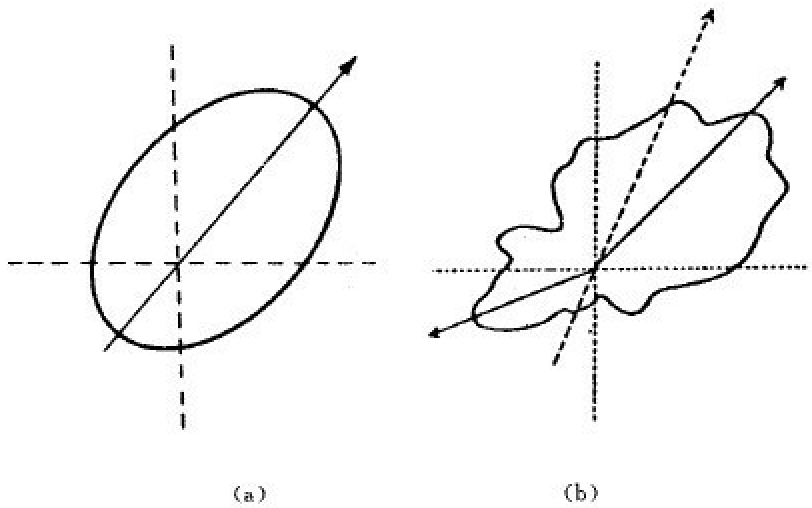


图1

爱丁顿对卡普坦的发现作了如下描述：

> 现在出版的“格罗宁根出版物”系列谈到这些问题，其中最有趣的当数第六本。但是，去图书馆查阅该书却是枉然，因为第六本中那些有趣之事并没有写出来。自然界发生了意外的变故，与第六本书原先详细叙述的设想不符。第六本书题为《宇宙速度之分布；第一部分理论》，研究的是当太阳的运动和距离变化时，恒星沿自身方向随机运动的统计特征。与此同时，还准备比较观察得到的奥维尔—
>
> 布拉德雷本征运动，以确定公式中的数值常数。这个理论尽管代表着当时公认的观点，可结果根本不对。甚至完全无法比较；人们不得不放弃使用这个公式。这就是卡普坦关于两个恒星流的伟大发现，它是1905年在南非举行的英国学术协会会议上宣布的。这个发现第一次揭示了恒星系统的某种结构，为研究相距遥远的恒星个体间的关系开创了一个新纪元。[8]

爱丁顿和卡普坦假设，可形式地将太阳周围的恒星看作两个群或者流，它们处于相对运动中，并且每个群内的运动是独立随机的。他们由此就可说明恒星本征运动分布的观察特征。这就是卡普坦爱—丁顿二星

流假说。根据这一假说，以前提出的麦克斯韦速度分布假设，也即

$$dN = N \frac{j^3}{\pi^{3/2}} e^{-j^2 |\boldsymbol{u}|^2} d\boldsymbol{u}, \quad ①$$

要用下式代替

$$dN = N_1 \frac{j_1^3}{\pi^{[1/2]}} e^{-j_1^2 |\boldsymbol{u}-\boldsymbol{u}_1|^2} d\boldsymbol{u} + N_2 \frac{j_2^3}{\pi^{[1/2]}} e^{-j_2^2} \left|\boldsymbol{u}-\boldsymbol{u}_2\right|^2 d\boldsymbol{u}, \quad ②$$

其中，N_1，N_2 分别为两个星流的恒星数目，u_1，u_2 是星流在局部静止系统中的速度，j_1，j_2 是星流中恒星相对平均速率。另外，速度分布（2）相对于局部静止系统，它满足：

$$N_1\boldsymbol{u}_1+N_2\boldsymbol{u}_2=0 \quad (3)$$

图2爱丁顿与卡普坦，1922年5月摄于罗马



图2爱丁顿与卡普坦，1922年5月摄于罗马

在格林尼治的那段时间里，爱丁顿写了好几篇论文讨论用分布（2）描述恒星运动的动力学。他还提出了确定星流N_1，N_2，u_1，u_2，j_1，j_2等参数的分析方法。根据所掌握的恒星本征运动方面的知识，爱丁顿求出了这些参数。爱丁顿的这些文章表现了理论与观测的非凡综合，反映了他对天文观测结果的犀利洞察力。

卡普坦—爱丁顿二星流假说颇令人满意，可卡尔·史瓦西就同样的观测给出了另外的解释。史瓦西的出发点在于，在适当取向的参照系中，用更一般的椭球分布代替麦克斯韦分布（1），即

$$dN = N\frac{j_1 j_2 j_3}{\pi^{3/2}}\exp\left(-j_1^2\boldsymbol{u}_1^2 - j_2^2\boldsymbol{u}_2^2 - j_3^2\boldsymbol{u}_3^2\right)d\boldsymbol{u}_1 d\boldsymbol{u}_2 d\boldsymbol{u}_3 \qquad ④$$

爱丁顿本人认为史瓦西公式是解释卡普坦发现最优美、最合适的方法。这种解释一直保留到现在。

爱丁顿1914年出版了第一本著作《恒星运动与宇宙的结构》。作为这一阶段成果的总结，爱丁顿书中以大量的篇幅对当时已有恒星运动知识，做了系统表述。但最后一章“恒星系统动力学”却开辟了有前途的新天地。爱丁顿先证明两星相遇不能有效改变单个恒星的运动方向，而后得出结论：决定恒星在六维相空间分布的函数f（x，y，z，u，v，w，t）需由恒星处于平滑引力势中的动力学轨道，也即六维刘维尔方程（Liouville equation，或我们现在所称无碰撞玻尔兹曼方程）的解来确定。

在1915年和1916年发表的论文中，爱丁顿求出了刘维尔方程的解，它与史瓦西的速度椭圆分布一致；特别是得到适合球对称星系的自洽解。与此类似，爱丁顿第一次提出了应用维里定理（Virial theorem）将群集恒星的平均动能与其平均势能联系起来的方法——至今这种方法在银河系和银河系族这样更大的天体情形下仍然有用。

因此，完全可以说爱丁顿创立了恒星动力学这一仍颇具生命力的学科。

Ⅳ

下面我要转向的话题无疑是爱丁顿对物理学最重要的贡献——创立现代理论天体物理学，开辟恒星结构、组织和演化的新学科。1916年为理解造父变星亮度的变化，激发了他对恒星结构的兴趣；随着《恒星的内部结构》的出版，这种兴趣达到了顶点。人们应当记住，在这同一个10年间，爱丁顿参与了对日食现象的观测，它提供了光通过引力场发生偏转的第一个证据——他后来称这件事为他在天文学研究中最激动人心的事件；除此以外，他还写出了《相对性的数学理论》（1923），更不用说《关于引力相对论的报告》（1915）以及他深受欢迎的两本书：《空间，时间与引力》（1920）和《恒星与原子》（1927）——这是硕果累累的10年。

在恒星内部结构领域，爱丁顿认识并建立了以下我们现在理解的基本原理：

1.辐射压对维持恒星平衡的重要作用，随恒星质量的不断增长越来越大。

2.恒星内部达到辐射平衡时的情况与对流平衡不同，辐射平衡时温度梯度由能源分布和恒星物质对优势辐射场的不透明度分布共同决定。确切地说有

$$\frac{dp_r}{dr}=-k\frac{L(r)}{4\pi cr^2}p,\ p_r=\frac{1}{3}\alpha T^4 \qquad (5)$$

和

$$L(r)=4\pi\int_i^r \varepsilon\rho r^2dr \qquad (6)$$

这里，p_r，κ，ε和ρ分别表示辐射压，恒星不透明度系数，每克恒星物质的产热率和密度。另外，α为斯蒂藩辐射常数，c为光速。

3.影响不透明度系数κ的主要物理过程取决于软X射线区的光电吸收

系数，也即取决于高度离子化原子的最内部的K层和L层的电离作用。

4.电子散射是星体不透明度的最终源泉，对于维持质量为给定值M的恒星来说，光度有一个上限。最大光度由下式给出：

$$L < \frac{4\pi cGM}{\sigma_e} \tag{7}$$

其中σ_e为汤姆逊散射系数，最大光度现在通常称作爱丁顿极限。爱丁顿极限在目前有关黑洞周围X射线源和吸积盘光度的研究中发挥着重要作用。

5.一级近似情况下，普通恒星（也即主序恒星）的质量、光度、有效温度关系对恒星能量源分布不是很敏感，因此，即使对恒星能量源了解不多，仍能建立某种关系与观测作比较。

6.氢变成氦是恒星最可能的能量来源。

爱丁顿60年前作出的这些推断，在当时以至现在仍然有效。

我想对其中的一些推断加以展开，以揭示爱丁顿应付这类问题的思路。

爱丁顿得出，辐射压是恒星平衡的一个因素，恒星质量越大则辐射作用就越大，我们首先考察确立这一结论的方法。我们可以回忆一下，爱丁顿带上所有的零写出太阳的质量时，他评论说，人们不应认为零的确切数目没有什么特别的意义，因为“大自然在乎”。

爱丁顿在《恒星的内部结构》一书中得出结论的方式是，他想象这样一位科学家：

在一颗云雾缭绕的行星上，这位物理学家对恒星闻所未闻。他计算着一系列大小不同气球体（globes of gas）的

辐射压与气压之比，比如说从质量为10克的球开始，然后是100克，1000克，继续下去，至第n个球，其质量为10n克。表1给出更有意义的部分结果。

表1

球序号	辐射压	气压
30	0.00000016	0.99999984
31	0.000016	0.999984
32	0.0016	0.9984
33	0.106	0.894
34	0.570	0.430
35	0.850	0.150
36	0.951	0.049
37	0.984	0.016
38	0.9951	0.0049
39	0.9984	0.0016
40	0.99951	0.00049

该表其余的部分主要是9和0组成的长串。只是第33号至35号球这一特殊的质量范围才有意思，这以后又变成9和0的串。如果说物质与以太（气压与辐射压）之间存在一场较量，这场较量完全是一边倒，只是在第33至35号球之间有些例外。可以预料那里出现了什么事。

“出现”的正是恒星。

我们的物理学家一直在云雾下面工作，现在掀开云雾的面纱，他得以仰视天空。那里发现有10亿个气球体，质

量几乎都在第33和35号球之间，也就是说质量在太阳质量的一

半至五倍之间。已知恒星中，最轻的恒星约为2×10^{32} 克，最重约为2×10^{35} 克。大部分在10^{33} ～10^{34} 之间，辐射压与气压抗争的严峻挑战就在这里开始了。[9]

表1的计算基于这样的假设：气压（p_g ）与辐射压（P_r ）之比β/（1-β）在整个恒星范围内为一常数；并且平均分子量μ为2.5。本来μ=1.0更切合实际一些，那样每个球的质量会以因子（2.5）^{2} =6.25的倍数增加。这个因子的意义并不大。可爱丁顿有两个重要问题没有涉及。其二，尽管计算结果清楚表明上述计算涉及质量大小和星体尺寸（带上所有的零！）的自然常数的组合，可爱丁顿并没有将它分离出来，鉴于他后来对自然常数的专注，这种疏忽颇令人惊讶。实际上，在我们感兴趣的范围内，决定星球质量的自然常数组合为

$$\left[\left(\frac{k}{H}\right)^4 R\,\frac{3}{\alpha}\right]^{1/2}\frac{1}{G^{3/2}} \qquad (8)$$

其中，H为质子质量，G为引力常数，k和α分别为玻尔兹曼常数和斯忒藩常数。斯忒藩常数的值为

$$\alpha=\frac{8\pi^5 k_4}{15h^3c^3}, \qquad ⑨$$

其中h为普朗克常数。将它代入（8）式，我们发现所涉质量大小的自然常数组合为

$$\left(\frac{hc}{G}\right)^{3/2}\frac{1}{H^2}\cong 29.2 \quad \cong 5.2\times 10^{34}\,g \tag{10}$$

可以这样说，现行的恒星结构和恒星演化理论的成功，很大程度上基于正确提供了上述恒星质量大小的自然常数组合（这里说明一下，质量大小的一般组合为

$$\left(\frac{hc}{G}\right)^{\alpha}\frac{1}{H^{2\alpha-1}}, \tag{11}$$

其中，α是任意的，α=1/2时，上式包含了普朗克质量[（hc/G）$^{1/2}$）]。

表1所示的计算中，爱丁顿没有提到的第二个问题是：为什么辐射压与引力抗争的范围与“恒星的发生”有关？在这个问题上，爱丁顿没有在他的“标准模型”（即辐射压与气压之比在整个恒星中为常数）上进行讨论，那样他本可以采用他自己的一条定理——稳定性定理，即，恒星中心的压力要比相同质量、相同中心密度、密度均匀物体的中心压力小，以此可以说明恒星中心的辐射压与总压力之比小于只考虑恒星质量所计算出的结果。这样一来表1由表2替代，而对被云雾围绕的行星上那位物理学家来说，结论是一样的。

表**2**

$(M/M_{\odot})\mu^2$	辐射压	气压
0.56	0.01	0.99
1.01	0.03	0.97
2.14	0.10	0.90
3.83	0.20	0.80
6.12	0.30	0.70
9.62	0.40	0.60
15.49	0.50	0.50
26.52	0.60	0.40
50.92	0.70	0.30
122.5	0.80	0.20
224.4	0.85	0.15
519.6	0.90	0.10

关于其他一些关键方面，有一项与恒星不透明性相关的工作，充分展示了爱丁顿进行天文学推理时，对待物理学理论的态度。在考虑有关恒星不透明性问题时，他请教了X射线及γ射线专家埃里斯（C.D.Ellis）并受益匪浅。另外，克拉默斯有一篇著名的论文，第一次对光电电离的原子截面从理论上作了估算，而这正是爱丁顿所需要的。不过，在克拉默斯论文发表之前，爱丁顿已经根据原子核直接俘获电子的假设提出了自己的理论，这种假设在物理学上是站不住脚的。他还坚持将他的理论与天文学联系起来，反对包括卢瑟福在内的许多物理学家提出的有说服力的证据。只是在克莱默斯论文发表并且其理论推断与实验结果被验证一致后，爱丁顿才放弃自己的理论。可根据克拉默斯的不透明性定律，天文观测结果与理论质光关系将相差十倍以上。1932年爱丁顿消除了这一偏差（斯特龙格伦也独立地做到这一点）。他的做法是采用氢和氦的高丰度值，那时罗素已通过对太阳大气元素丰度的分析确定了这一点。然而由此得出恒星的构成并不都是一样的结论，爱丁顿过很久才承认。事实上，爱丁顿更早的时候就认识到，只要假设氢丰度很高，并且恒星

平均分子量接近于1，就可以解决不透明性偏差问题。但这一假设会破坏自己关于“恒星的发生”的证据。按照爱丁顿的特点，他本该得出这样的结论：“我宁可给这种不一致性找到另外的解释。”

爱丁顿关于恒星能源的预测大概是他作出的最壮观的预言。1920年8月24日他在卡迪夫给英国学术协会作的报告中，包含了所有天文学文献中最富预见性的论述。

> 只是出于传统的惰性才使得有些人仍然不肯放弃收缩假说，其实，它不再活着，只是一具未被埋葬的尸首。可是，在我们决定埋葬这具死尸之前，我们得坦率地认清我们目前所处的位置。恒星以未被我们了解的方式，动用某个巨大的能量储备库作能量源。这个能量库几乎只可能是亚原子能量，众所周知，这种能量在所有物质中都大量存在。我们有时梦想总有一天人类学会如何让它们释放出来，为己所用。这一储备几乎取之不竭，只要它能被发掘出来。太阳的能量足够它的热量继续输出1150亿年……
>
> 阿斯顿进一步确证氦原子的质量比进入其中的四个氢原子的质量之和还小——至于这一点，化学家们无论如何是同意的。聚合过程中质量损失总计1/120，氢的原子量为1.008，氦的原子量只有4。我不想详述阿斯顿给出的精彩证明，因为你们可从他那里了解这些。质量不可能消灭，其赤字部分只可能在嬗变过程中以能量的形武释放出来。于是我们立即可以算出氢变为氦时释放的能量。如果最初恒星质量的5%由氢原子组成，这些氢原子逐渐合成更复杂的元素，其间释放的能量就超出我们的需求，我们也就无需进一步寻找恒星能量的来源。
>
> 确实，如果恒星的亚原子能量可随意用来维持其巨大熔炉，控制这一潜在能量造福人类——或毁灭人类的梦想，就似乎离现实更近一步。[10]

在关于恒星能量来源问题上，老一代的天体物理学家会回忆起爱丁顿在进行辩驳时的一段名言：

> 举例来说，已经有人表示反对，认为恒星温度不足以使氢嬗变为氦，从而对这一可能的能源不予考虑。可氦是存在着的，对于批

评家来说，力陈恒星温度如何不够并无太大用处，除非他能够向我们指出一个更热的地方来。[11]

V

我在开头曾说过，爱丁顿对恒星内部结构的兴趣，源于他试图解释造父变星所显示出的恒星可变性和周光关系而做出的努力。里特尔（Ritter），早些时候对处于对流平衡中的气态恒星绝热脉动进行了分析，爱丁顿对分析结果做了推广。并将它应用在按他的标准模型建立的辐射平衡恒星上。爱丁顿接着将得到的周期公式与他的质光关系结合起来，从而能够以一种普遍的方式解释观测得到的造父变星的周光关系。有关恒星可变性的脉动理论就这样逐渐建立起来。

爱丁顿对造父变星可变性的初步分析，没有提供诸如恒星亮度、有效温度和径向速度等变量间的标准位相关系。然而他清楚认识到，大量的元素氢和氦在恒星外层被电离并形成对流区，只有仔细研究那里的能量转移机制，才能理解这些位相关系。爱丁顿以后几年多次回到这一问题。事实上，他最后发表的论文之一就是专门讨论这个问题的。只是到了后来，由于史瓦西、莱托（P.Ledoux）和克里斯蒂（R.christy）的共同研究，才最终解决了这个问题。

虽然爱丁顿对天体物理学的主要贡献在恒星结构领域，这决不是说他对天体物理学其他领域的贡献不重要。他提出了辐射转移问题的一种近似解决方法——“爱丁顿近似”；他解决了恒星大气中谱线的形成问题，这在恒星大气理论发展初期常常用到；他还讨论了相距很近的双星系统的反射效应，人们根据双星的星蚀分析光线弯曲，以确定单个恒星的质量时，这种效应必须考虑。爱丁顿在后一项工作中考虑的问题。正是分层平面大气中有关光的散射与透射这一更大问题的原型。这一学科后来才逐渐成熟起来。

在天体物理学的这些研究领域中。最重要的或许要数爱丁顿引入了“稀释因子”（dilution　factot）——他发明的这个词沿用至今，在确定星际空间电离状态时，这个因子可以使我们考虑到主辐射场场强的减弱。也是爱丁顿最先修正了“生长曲线”方法，将它应用于星际吸收谱线问题。安索尔德（A.Unsold）和明纳尔特（M.Minnaert）提出的“生长曲线”方法，可用于根据恒星吸收谱线强度得出元素的相对丰度。

爱丁顿对星系动力学和天体物理学的兴趣集中体现在他以下预言上：由星际吸收谱线确定的径向速度与银河纬度有一定关系，由这种关系确定的径向速度存在一个幅度，这个幅度应是恒星吸收线所表现出的一半；后来斯特路维（O.Struve）和普拉斯克特（J.S. Plhaskett）通过观测完美地证实了这一预测。

至于《恒基的内部结构》，它包含了目前为止我已介绍过的大部分内容，罗素曾说过："这本著作堪称一部第一流的杰作。"[12]

Ⅵ

我在叙述爱丁顿对恒星结构的贡献时，只字未提他先后与琼斯（Jeans）和米尔恩进行的连续论战。从现在流行的观点看，当时争论的问题似乎并不是很重要：随着对恒星能源的理解加深，一些悬而未决的问题需要不同的表述和不同的解决办法。应该说，爱丁顿对待自己科学对手的方式并不总是很公平，在1929年12月的皇家天文学会会议上，米尔恩宣读了一篇论文。爱丁顿对此回应如下：

> 米尔恩教授没有详细谈及为什么他的结论与我的有如此天壤之别。我对文章其余部分兴趣索然，若对它的正确性还存有丝毫幻想的话，那太荒谬了。[13]

争论时使用的语言有时显得近乎刻薄，以下摘自琼斯发表在《天文台》杂志的两封信，充分说明了这一点：

> 人们在等温平衡方面已经做过大量的工作，因此很难理解爱丁顿教授怎会抱有这样的错觉，似乎他正在未被探索的领域从事开拓性的工作。他完全不参考其他人的理论工作（除了引用埃姆登的一些数值计算结果外），这一事实表明情况确实如此。（1926年8月）

> 结束本信时请允许我谨向爱丁顿教授保证，如果在将来他能够戒除对我的工作的野蛮攻击，尽管他无法对我的工作加以证实，同时一旦发现我以前的工作对他有所裨益而按惯例表示他的谢意，从而清除我们之间长期不和的根源的话，我会感到莫大的快乐。我更看重的是上述请求的后一部分，因为我发现我引进天体物理学的一

些最富成效的思想，如物质的湮灭是恒星能量的来源，以及高度电离的原子和自由电子是组成恒星的物质等，现在都相当普遍地归于爱丁顿教授了。（1926年11月）[14]

我想以一段轻松的轶事中止这些让人不愉快的插曲。

大家知道，爱下顿有时喜欢看赛马，不时带他的姐姐一起去参观纽马克赛马大会。哈代一定知道此事，因为有一次我听他问爱丁顿是否赌过马。爱丁顿承认曾经赌过，“不过只有一次。”哈代很想知道那次赛马的情况；爱丁顿解释说一匹名叫琼斯的马飞跑着，他无法抗拒诱惑而在它身上下了注。问他赢了没有，爱丁顿带着他特有的微笑回答说：“没有！”

（2）爱丁顿：广义相对论的阐述者和倡导者

我上一次讲座主要探讨了爱丁顿对理论天体物理学的贡献，并具体说明罗素对他的评价是恰如其分的。罗素认为爱丁顿是他那个时代天体物理学最杰出的代表。我在这次讲座里准备涉及这样一些内容：爱丁顿是广义相对论的阐述者和倡导者；为验证爱因斯坦关于光在引力场中发生偏转的预测，格林尼治—剑桥联合考察队观测了1919年5月29日的日食现象，爱丁顿在这次考察中起到什么作用；在长达16年以上的时间里，爱丁顿在宇宙学以及——用他自己的话说——“统一量子论和相对论”方面所作的努力。不过，与上次讲座相比，这次讲座恐怕并不完全令人愉快。

我想还是先从较愉快的一面谈起。

Ⅶ

爱因斯坦1905年创立了狭义相对论原理，接下来的十年里他致力于使牛顿的引力理论与狭义相对论中的两个原理，特别是任何信号都不能以超过光速的速度传播这一原理相一致。在经历多次失败后，1915年夏秋季爱因斯坦在写给柏林科学院的一系列引人注目的短信中，达到了目标。战争仍在进行，要不是荷兰的中立以及爱因斯坦与洛伦兹（H.A.Lorentz）、埃伦菲斯特和德西特（W.de Sitter）的私交（见图3），爱因斯坦成功的消息就不可能越过英吉利海峡（更不必说大西洋

了）。德西特将爱因斯坦论文的复本寄给爱丁顿，并在1916—1917年间将他自己的三篇论文送交皇家天文学会，这三篇文章部分是阐述性的，部分是他本人的贡献。现在所称的德西特宇宙，在上述三篇文章的最后一篇里做了描述。

那时爱丁顿担任皇家天文学会秘书职务，他必须对德西特的论文进行处理。从他在1917年12月的皇家天文学会议上对德西特三篇文章最后一篇所做的说明，人们可以推测他仔细阅读了这些文章，并亲自进行了审定。[15]

人们会记得，爱因斯坦在最后一封信中阐述了他的基本场方程，结束时还预言般地说："任何充分理解这一理论的人，都无法逃避它的魔力。"爱丁顿一定是着了魔，因为不到两年时间，他就为伦敦物理学会写出了《关于引力相对论的报告》。这是一篇情绪激昂的报告，写得清晰而又简洁，即使在今天，它仍是低年级学生的一本好的入门读物。

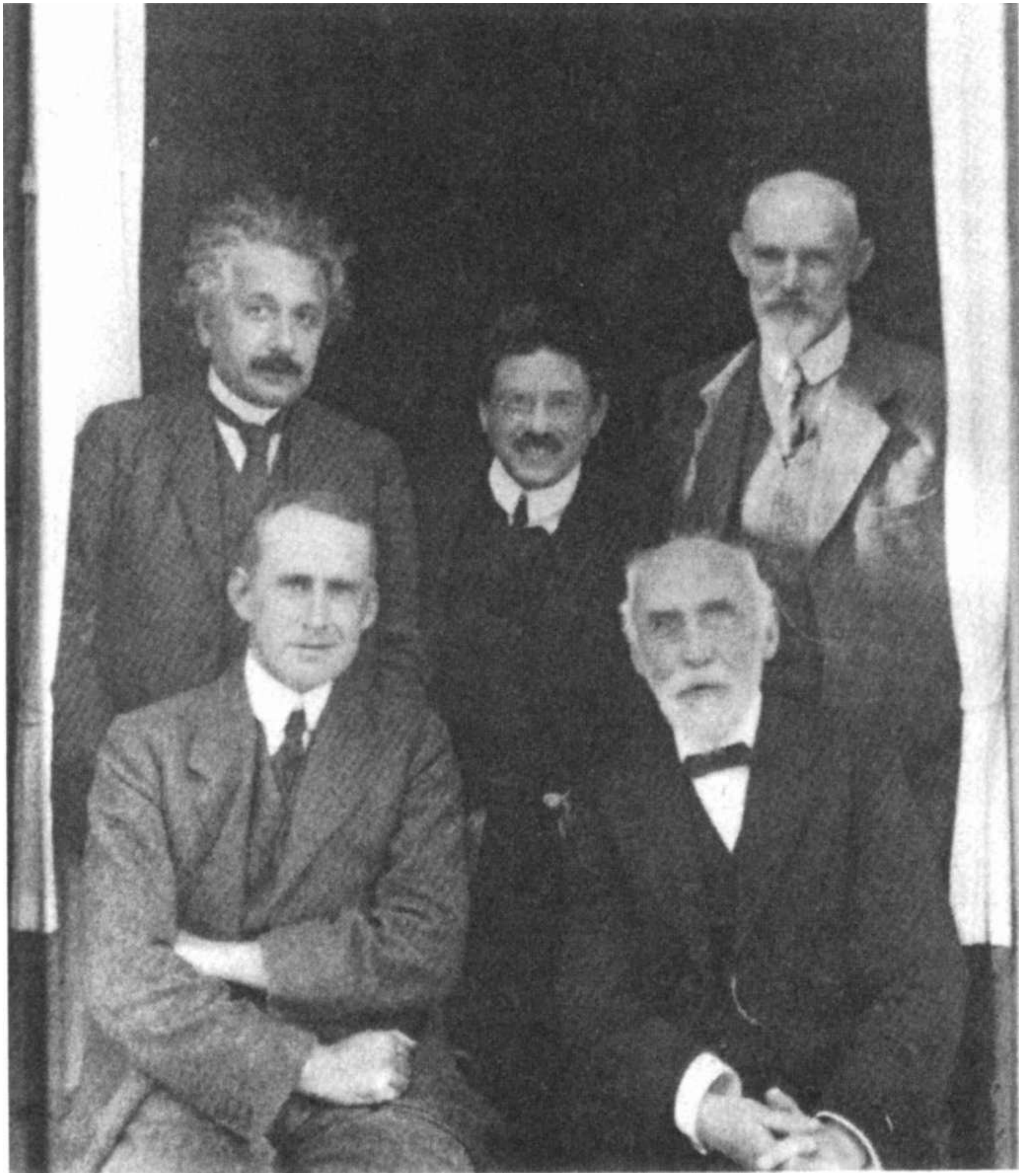


图3（左上起顺时针）A.爱因斯坦，P.埃伦费斯特，W.德西特，H.A.洛伦兹和A.S.爱丁顿。1923年9月26日由德西特的大儿子（现为莱顿大学地质学教授）摄于德西特书房

Ⅷ

毫无疑问，正是爱丁顿对广义相对论的热情，成功地影响了他的亲

密朋友和同事费兰克·戴逊爵士，并使这位皇家天文学会会员对广义相对论着了迷；不久，他们共同制订了一个考察计划，如果“那个时刻到来之际文明状态允许的话”（引自琼斯之语），他们打算对1919年5月29日的日食现象进行观测。对于自己参与制订这次考察计划并取得成功一事，爱丁顿认为是“我在天文学研究中最激动人心的事件。”有许多有趣的事情，真不知该从何说起。请让我先谈谈爱丁顿曾给我做的说明。

记得有一次，我告诉爱丁顿说我非常钦佩他的科学敏感性：在未来显得异常暗淡的情况下，他还制订了那次远征的计划。令我吃惊的是，爱丁顿对这一评价未做任何赞许，并还告诉我，如果全由他自己决定，他决不会组织那样的考察，因为他完全相信广义相对论的正确性！他还介绍了参与这次考察的经过。这些我在《皇家学会评论和记录》[16] 上介绍过，不过请允许我再重复一下。

1917年，战争已进行两年多，英国制订了对所有强壮男子的征兵法令，爱丁顿时年34岁，符合征兵条件，但作为热心虔诚的教友派信徒，他是一个基于道义的反战者。人们知道而且预料他会以此为理由要求延缓服兵役。第一次世界大战期间，英国国内的舆论对于那些真诚的反战者是非常不利的，事实上，与拒服兵役者结交被认为是一种社会耻辱。那个时候，剑桥的一些忠实朋友们，如约瑟夫·拉莫爵士（Joseph Larnlor，拉莫进动即以他的名字命名）、纽沃尔（H.F.Newall）教授及其他一些人，试图通过内政部以使爱丁顿缓服兵役，他们提出的理由是：爱丁顿是一位著名科学家，让他在军队服役不符合英国的长远利益。英国科学家们对莫斯利（Moseley）阵亡加里波利的事记忆犹新。拉莫等人的努力差不多就要成功，内政部给爱丁顿去了一封信，他只需签上自己的名字寄回就行了。可爱丁顿在回信上加了一个附注，大意是说如果不能因为所述理由缓服兵役的话，他就以基于道义的反战来要求缓服兵役。这样的附注自然将内政部置于窘境，因为规定凡拒服兵役者必须送入军营；拉莫他们为此也非常生气。但爱丁顿告诉我说，他不明白他们生气的理由。当他表达自己意见的时候，他的许多教友派朋友正在北英格兰的军营里削马铃薯皮，他认为他没有任何理由不和他们一起。显然是由于戴逊的调停（作为皇家天文学家，他与海军军部联系密切），爱丁顿终于延缓了兵役，他们达成明确的条件：如果战争在1919年5月前结束，爱丁顿就要率领一个考察队去验证爱因斯坦的预言！

关于这些事件已公开发表了一些略有不同的说法；不过它们仅在侧

重点和较次要的情节上有所不同。好在爱丁顿本人也曾叙述了这次考察的计划和进行过程。他写道：

> 光线的弯曲影响太阳附近的可见恒星，因而只有在日全食期间，月亮遮住眩目的太阳光线时，才能进行观测。即使在那时，仍有大量的日冕光线在日面上延伸很远。因此，必须观测太阳附近很亮的恒星，它们才不会淹没在日冕的强烈照射之中。另外，这些恒星的位移只能相对于其他一些恒星来测量，那些恒星离太阳较远而且位移较小；我们需要一定数量的外围亮星作为参考点。
>
> 在迷信年代，自然哲学家若想进行一项重要实验，就会求助占星家确定实验的黄道吉日。今天的天文学家借助更可靠的理性，经研究后人们发现，一年中观测光线最好的日子是5月29日。理由是太阳绕着黄道的周年转动中，要穿过稀密不同的恒星域，但只是在5月29日这一天，太阳处于一团由明亮恒星——毕宿星团的一部分——组成的特殊区域的中央。这真是理想的最佳星域。如果在历史上别的时候提出这样的问题的话，或许得等上好几千年才会在这个幸运日子发生日全食，可运气好得出奇，就在1919年5月29日发生了日食……
>
> 1917年3月，皇家天文学家（戴逊爵士）提醒人们注意这个难得的机会；皇家学会和皇家天文学会组成的一个委员会开始了观测的准备工作……[17]
>
> ……在大战中的1918年就开始制订计划，但直到最后一刻还不能确定考察队能否出征……戴逊爵士在格林尼治组织了两个考察队，一个去巴西的索布拉尔，另一个去西非的普林西比岛。克罗梅林（A.C.D.Crommelin）博士和戴维逊（C.Davidson）先生去索布拉尔，柯丁罕（E.T.
>
> Cottingham）先生及我去普林西比。
>
> 仪器制造商们在停战前无法动工；同时因考察队必须在二月份起航，准备工作就相当匆忙。去巴西的那一组碰上好天气；由于一些偶然情况，他们的观测结果一直拖了好几个月才整理出来，不过最终还是提供了最具决定性的证明。我在去普林西比的那一组。发

生日食那天，那里正巧赶上阴雨和多云天气，几乎没有任何希望。将近全食的时候，太阳开始发暗；我们仍然完成了工作计划，但愿情况不会太糟糕。云层一定是在日全食结束前变薄了，因为尽管有许多照片报废，我们还是得到两张底片，上面有我们所期望的恒星图像。将这些照片与对同样星域在太阳处于别处时拍下的照片做比较，两者之间的差别就表明，恒星由于光线经过太阳附近发生弯曲而造成表观位移。

当时我们面临的情况，存在三种可能性；或许根本不存在什么偏移，也就是说光线不受引力的影响；或许存在“半偏移”，正如牛顿所言，光受重力的影响，遵从简单的牛顿定律；或许存在“全偏移”，从而证实了爱因斯坦定律而非牛顿定律。我记得戴逊将所有这些解释给我的同事柯丁罕听了以后，将主要思想总结为偏差越大，就越是激动人心。柯丁罕问：“如果我们得到双倍偏移，那到底意味着什么？”“那么，”戴逊回答说，“爱丁顿就会发疯，你只得孤身一人回去。”

事先就安排当场对底片进行测量，这倒不全是因为急不可待，而是担心回家途中发生什么意外，所以我们立即检查了一张成功的底片。天文测量中所得到的数值与理论预言的一样大，所以一张底片几乎就可决定问题的结果，尽管肯定还会通过其他的底片来加以证实。日食三天之后，计算最终完成，我知道爱因斯坦理论经受住了考验。科学思想的新观点取得了胜利。柯丁罕用不着孤身一人回家。[18]

1919年11月6日，皇家学会主席汤姆逊爵士（J.J.Thomson）主持了皇家学会和皇家天文学会联合会议，会上报告了这次考察的科学结果。这次会议引起了少有的关注，我清楚记得卢瑟福有一次回忆起这次会议受关注的情况。

1933年圣诞假期间，在三一大厅举行的餐会之后，我与卢瑟福、爱丁顿、帕特里克·杜瓦尔（Patrick Duvat，著名几何学家）以及阿莫斯爵士（Maurice Amos，20世纪20年代曾任埃及政府高级司法顾问）围坐在教员休息室的火炉旁交谈着。谈话中间，阿莫斯有一次转身朝向卢瑟福说：“我不明白为什么爱因斯坦会比你赢得更高的赞誉，毕竟是你发明了原子的核模型；并且这一模型给今天物理学的一切奠定了基础，它的

应用甚至比牛顿的引力定律更普遍。还有，爱因斯坦的预言与牛顿理论的差别微乎其微，我不明白所有这些纷纷扬扬的小题大做到底是为什么。”听罢这席话，卢瑟福转向爱丁顿说：“你对爱因斯坦的名望负有责任。”他态度严肃地说：

> 战争已经结束，维多利亚和爱德华时代的自满已被粉碎。人们觉得一切价值和理想都迷失了方向。突然间，他们得知一位德国科学家提出的一个天文学预言被英国天文学家们证实，这些英国科学家特意去巴西和西非进行了考察，并且还在战争期间就对这次考察做了准备。天文学从来就吸引着大众的想象力；一项天文学发现，超越了世俗的冲突，拨响了人们敏感的心弦。在皇家学会的会议上报告这次英国考察队的结果，这在所有的英国报刊上都成了标题新闻，宣传的飓风跨过了大西洋。从此以后，美国新闻界把爱因斯坦捧上了天。

结束这一节报告之际，我想引述琼斯向戴逊颁发皇家天文学会金质奖章时说过的一段话：

> 1918年，战争中最黑暗的日子里，格林尼治天文台和剑桥大学分别组织考察队；打算对1919年5月的日食进行观测，如果那个时刻到来之际文明状态允许的话。这是对爱因斯坦广义相对论的决定性考验。1918年11月签署停战协定后不久，考察队出发了，并且后来带回的结果彻底改变了天文学家关于引力本质的概念以及普通人关于自己生活在其中的宇宙的本质概念。如果这种成功的荣誉必须在戴逊爵士和爱丁顿教授之间分开的话，坦率地讲，我真不知道该以什么比例作出分配。依我之见，与其说是平分荣誉，不如将全部荣誉归于每个人，因为如果他们中的任何人没有发挥作用，或者是缺乏洞察力、热心，或者缺乏把握时机的本领，我怀疑这次考察还能否完成，最早通过观测确定时间和空间到底为何物这一伟大的荣誉，很可能落入他人之手。[19]

IX

我想对这件事作些补充。

戴逊和爱丁顿将他们考察结果的报告讲完后，汤姆逊从会议主持人

角度作了如下评述：

> 事实上，牛顿在其《光学》中的第一个问题提出了这一点，根据他的假定本可以得出现有光弯曲数值的一个半值来。但是，这次测出的这个结果并不是孤立的，它是科学思想整体的一部分，影响着物理学最基本的概念……它是牛顿时代以来引力理论方面得到的最重要的结果，因而应当在和牛顿密切相关的皇家学会的会议上宣布，这样做是十分恰当的……
>
> 如果爱因斯坦的理论是对的，我们对引力就得采取一种全新的观点。如果能够证实爱因斯坦推理的正确性——它已经经受了与水星近日点和现在这次日食有关的两次严峻考验——那么他就是人类思想的一项最高成就。这一理论的弱点在于表述上的巨大困难。无论是谁，如果对不变量理论和变量微积分学缺乏透彻的了解，就不可能理解新的引力理论。[20]

正如汤姆逊提到的，广义相对论理解上的“困难”，当时并且以后很长一段时间一直也是公认的看法。确实，不久就有“世界上只有三个人懂广义相对论”的神话。事实上这一神话就产生于这同一次会议。

爱丁顿（在我前面提到的那次三一学院餐后谈话中）回忆起，就在皇家学会和皇家天文学会联合会议散会时，希尔伯斯坦（Ludwig Silberstein）走到他跟前说：“爱丁顿教授，您一定是世界上三个懂广义相对论的人之一。”正当爱丁顿对这一说法犹豫不决时，希尔伯斯坦又说：“不要谦虚，爱丁顿先生。”结果，爱丁顿回答说：“正好相反，我在考虑第三个人会是谁。”

我想顺便说一下，人们极大地夸张了理解广义相对论的这种公认的困难，这促成几十年来这一学科的停滞。许多在60和70年代取得的进展，本可以在二三十年代就很容易地得到。

爱丁顿很喜欢重复戴逊对柯丁罕说的话：“爱丁顿就会发疯，你只得孤身一人回去。”在1932年1月的一次皇家天文学会会议上，当佛伦里奇（Finlay Freundlich）报告说他的日食观测结果给出的光偏移值远远大于爱因斯坦预言值时[21]，爱丁顿重复了戴逊的话，气势咄咄逼人！但里克天文台在1923年4月的皇家天文学会会议上报告他们1922年观测得

出一致的结果时，爱丁顿作了如下评论：

> 我想是《怪兽搜捕记》[18]中的贝尔曼订下“事说三遍即为真”的规矩。现在恒星已经对三支独立的考察队重复三次，所以我确信结果是对的。[22]

最后，我想谈一个概率论问题，它在30年代末简直有些声名狼藉，最初就产生于格林尼治—剑桥考察队。问题本身以及爱丁顿解决问题的方式，显示出他对概率论的深刻理解。在星流研究中，爱丁顿必须对大量的不确定观测结果进行综合处理，他于是对概率论产生了兴趣（见《观测的综合》）。

你们会记得，两支日食考察队分别由克罗梅林、戴维逊（他们去索布拉尔）以及柯丁罕、爱丁顿（他们去普林西比）负责，在考察队出发前的一次餐后谈话中，克罗梅林暗示可能会出现下述情况：

如果C、C′、D、E各自独立地在三次讲话中说一次真话，D宣称E是一个说谎者，C′否定了D的话，C又肯定了C′的话，那么E讲真话的概率是多少？

爱丁顿在其《科学的新途径》。[23]中叙述这一问题时，将C、C′、D、E换成了A、B、C、D，并给出答案为25/71。他后来说，“我轻而易举给出答案时，没有料想到我会收到那么多信件。”许多人，包括丁戈尔（Dingle）在内，都认为问题的阐述含糊，爱丁顿给出的答案决非最显而易见的。爱丁顿觉得，这个问题明确告诉我们（任何一个通情达理的人都会同意），它作了2条陈述：[24]

（1）D作了一个判断，比如说，X；

（2）A作的判断为“B否认C反驳了X”。

需要求得的就是X为真的概率。爱丁顿将他的解答解释如下：

> 我们不知道B和C作了任何相关的判断。比如说，如果B真实地否定C反驳了X，我们也毫无理由认为C肯定X。
>
> 我们可以找出与已知条件不一致的组合仅为：

（α）A说真话，B说假话，C说真话，D说真话，

（β）A说真话，B说假话，C说假话，D说假话。

因为如果A说假话，我们就不知道B说什么；如果A和B都说真话。就不知道C说什么。

既然（α）和（β）在81次机会中分别出现2次和8次，D的27次真话和54次假话就减为25次和46次。所以，所求概率为25/71。[25]

爱丁顿的答案无疑是对的，尽管仍然有人对他持有异议。

X

我现在谈谈爱丁顿对经典相对论的其他方面贡献。我认为爱丁顿对广义相对论最伟大的贡献，在于他在《相对论的数学理论》中对该课题令人惊奇的处理方法。我经常用到它。此外，在他的数学处理中还点缀着许多引人注目的格言。我最喜欢的一句是："空间不是大量密集在一起的点；而是大量互相联结在一起的距离。"[26]

下面这段话摘自他的"对手"琼斯对爱丁顿这本书所做的评论，它很好地概括了其显著特点：

> 我们处处都可以感受到辛勤刻苦的劳动和一丝不苟的审慎；一章一章地往下看，每次都觉得对问题的说明总是再好不过的。由于作者为之付出的心血，数学家读这书会轻松而又愉快……
>
> 本书的文体通篇清晰易懂，令人敬佩；它完全达到了我们期望的爱丁顿教授的高水平，这样的赞誉是再恰当不过的。[27]

除此以外，我想简要介绍一下爱丁顿三项具体的研究，它们清楚表明他对经典相对论的领悟和理解。

首先，现代大多数相对论研究者都熟悉这样一个事实：在今天所称为事件视界（event　horizon）的地方。史瓦西度规的表观奇异性乃是选择坐标系的结果，没有任何更进一步的意义。爱丁顿在1924年明确地给

出一种变换——现在称作爱丁顿芬—克尔斯坦变换，阐明了这一事实。[28] 然而，应当说爱丁顿是因为别的目的得到这一变换的，并且从当时的情况看不出他正忙于坐标奇异性的问题。

第二，爱因斯坦用辐射源的不定四极矩表示引力辐射发射率，目前对爱因斯坦的原始公式正确与否有争论，所以回顾一下历史是很有意思的。早在1922年，爱丁顿就对刚性自旋杆的引力能量发射率作了清晰完整的演算，并得出正确的结果，他还顺便发现爱因斯坦原始公式里错了一个因子2。[29]

第三，爱丁顿和克拉克在1938年发表的一篇研究报告中，独立于爱因斯坦、英费尔德（Infeld）及霍夫曼（Hoffman），研究了广义相对论中的N体问题，并求出了我们现在称作一阶后牛顿近似中的度规系数。[30] 然而，爱丁顿和克拉克既没有将问题转换为哈密尔顿形式，也没有得出与十个运动方程积分相类似的经典结果。他们主要关心质心的运动。给出他们所述问题的一般性解答，需要始终在一阶后牛顿近似内，得出处于匀速相对运动中的两个参照系之间的坐标变换，也即后伽利略变换。爱丁顿和克拉克没有得到这样的变换；但他们确实解决了他们想要解决的问题，即两质点相对运动中的一种限制性问题。

前面所举的例子，特别是最后一个例子说明，如有兴趣的话，爱丁顿能够解决经典相对论中很艰深、很复杂的问题。可他似乎兴趣不大。

XI

对于自己在经典相对论方面的贡献，爱丁顿最看重他对外尔理论的推广，该理论试图将引力理论与电磁学统一起来。1954年，有人在爱丁顿论文中发现了一份文件[31] ，他在这里泛泛谈到他认为是自己主要科学成就的那些内容，其中突出提到推广了外尔的理论以及“与此相联系，他对引力定律的解释”，即$G_{\mu\nu}=\Lambda g_{\mu\nu}$，其中Λ为宇宙常数。无论如何，他当时所形成的心态，在以后一些年中越发明确地成为他永恒的基石。因此，我打算简明扼要地阐述一下外尔理论及爱丁顿所作推广的本质。

在爱因斯坦系统表述他的广义相对论时，人们认为整个物理世界中只需它给出两种场的描述：引力场和电磁场。爱因斯坦已经说明如何将

引力场并入时空结构，下一步很自然就是将电磁场也并入时空结构。显然，要完成这样的归并，必须适当推广黎曼几何，拓广爱因斯坦理论的几何基础。为了完成这种推广，外尔和爱丁顿考虑了沿着一个封闭的无穷小环路平移一个向量的最后效果。在黎曼几何中，向量描画出这样一个封闭环路后会改变其方向，但长度不变。外尔假定长度也会改变一个与其初始长度成比例的值；爱丁顿则（首先）假定长度的改变是任意的。

在外尔理论中，根据其基本假设，黎曼几何的克里斯托菲尔（Christoffel）联络Γ_{ij}，k变成

$$\Gamma^{*}_{ij,k} = \Gamma_{ij,k} + \frac{1}{2}\left(g_{ik}\varphi_j + g_{ik}\varphi_i - g_{ik}\varphi_k\right), \tag{12}$$

其中，φ_i （i=1，2，3，4）是一些平滑函数。另外，在外尔理论中，我们要求所有的几何关系和物理定律不仅对任意坐标变换不变（如同在爱因斯坦理论中一样），而且对度规变换，即变换

$$\varphi_i \to \varphi_i - \frac{1}{\lambda}\frac{ax}{ax^i}, \tag{13}$$

也应不变，其中λ为一任意函数。根据这些假设，外尔证明

$$F_{ik} = \frac{a\varphi_k}{ax^i} - \frac{a\varphi_i}{ax^k}, \tag{14}$$

具有麦克斯韦张量的所有性质；他还通过这种方法成功达到引力与电磁学的统一。

对引力理论来说，外尔理论最重要的结果就是在没有电磁场的情况下，爱因斯坦方程（真空情况下）

$$G_{ij}=0 \tag{15}$$

变成

$$G_{ij}=\Lambda g_{ij} \tag{16}$$

这里，Λ是一个普适常数，g_{ij} 和G_{ij} 分别是度规张量和爱因斯坦张量。方程（16）中的常数Λ就是爱因斯坦早在1917年就引入的宇宙常数，现在想起来，引入该常数的目的是为了使他的理论接受一种静止、均匀和各向同性的宇宙模型。

爱丁顿对外尔理论的推广相当于以下式代替方程（12）：

$$\Gamma_{ij,\ k}=\Gamma_{ij,\ k}+K_{ik,\ j}+K_{ik,\ i}-K_{ij,\ k}, \tag{17}$$

其中，Kik，j是某个秩为3的协变张量（暂未详细说明）。另外，在没有电磁场的情况下，我们就会得到包含宇宙常数项的爱因斯坦方程。

根据以上方程，自然会得出宇宙常数项，这一事实使爱丁顿确信在爱因斯坦方程中包含该项的必要性；并且这一点成为他的观点的核心。正如他自己解释的：

> 在任意一点上沿任意方向的曲率半径都与处于同一点和取同一方向的特定物质单位的长度之比为一常数。[32]

或者反过来说，

特定物质结构的长度与它所处位置及所取方向的宇宙曲率半径之比为一常数。

爱丁顿通过各种不同的方式表达他的这一核心思想。因此他说：

我们知道，无论物质结构与它周围的真空达到平衡时所遵从的实际规律是什么，爱因斯坦引力定律是用物质测量工具来测量世界时，几乎不可避免的结果。[33]

还有，

电子不会知道它应该是多大，除非空间存在独立的长度以便电子能用来度量它自己。[34]

确实，爱丁顿认为回复到不带Λ项的爱因斯坦方程相当于回复到牛顿理论：

我认为回复到牛顿理论与去掉宇宙常数是一回事。[35]

他的这种绝对信念导致产生下面这类极端性的说法：

令Λ=0就意味着退回到不完全相对论——这一步只能认为是退回到牛顿理论。[36]

……在我看来，宇宙常数的地位不可动摇；即使相对论名声扫地，宇宙常数将是最后坍塌的堡垒。去掉宇宙常数会动摇空间的根基。[37]

不过，并非只是爱丁顿持有这一观点。20世纪50年代末，有一次我曾询问勒迈特（Lemaitre）：以他所见，在我们的基本物理概念中广义相对论带来的最重要变化是什么。他毫不迟疑地回答是“宇宙常数的引入！”同样，爱因斯坦在1923年写给玻尔（Niels Bohr）的一封信中明确表示，“爱丁顿比外尔更接近真理。”[38] 确实，爱因斯坦提出的最后一种“统一场论”（薛定谔也独立地完成了该项工作），与爱丁顿对外尔理论的推广有很多共同之处。但外尔则有不同的看法，1953年他写道：

至于说爱丁顿本人对这一理论的创造性贡献，我认为主要包括

> 两方面的内容。首先是他关于仿射场理论（affine field theory）的思想，其次是他后来力图用认识论原因解释那些似乎参与了宇宙构成的纯数……
>
> 他的第一项贡献显然已经取得了成果。当爱因斯坦为这种理论阐述其作用原理时采用了它（我错误地相信爱丁顿当时认为这一理论是不必要的）……可我也对爱因斯坦最近的统一场论相当怀疑。我坚信关于引力本质的最后定论还未做出，并且我觉得解决这一问题的方向与爱丁顿和爱因斯坦最后的思想颇不相同。最终解决这一问题还需等待一段相当长的时间。[39]

尽管爱丁顿态度很肯定，可宇宙常数后来的命运一直是盛衰多变。当人们发现弗里德曼宇宙模型为解释哈勃膨胀这一简单事实提供了令人满意的根据时，爱因斯坦和德西特在一篇合写的文章中表示，没有宇宙常数也可以。鉴于对“撤销”Λ有许多夸大其辞的说法，准确记录下他们当时说过的话很有意义：

> 从历史上看，在场方程中引入“宇宙常数”项Λ是为了使我们能够从理论上解释静止宇宙中有限平均密度的存在。现在看来，不引入Λ项也能达到同样的结果……
>
> ……然而，曲率（常数Λ）必然是可以确定的，而且随着观测数据精度的提高，我们将会确定其符号及数值大小。[40]

爱丁顿曾写到爱因斯坦和德西特文章发表后不久他与爱因斯坦的会面，以及德西特写给他的一封信。它们就这个问题从侧面作出了有趣的说明，并且对爱因斯坦撤回Λ一事的极端性说法提出了疑问。

> 不久，爱因斯坦来我处探访，我在这件事上责备了他。他回答说：“我自己认为这篇文章不是很重要，可德西特却认为很好。”爱因斯坦刚走，德西特就写信说要来访。他还说：“你将会看到我和爱因斯坦的那篇文章。我认为结果价值不大，可爱因斯坦似乎认为它很重要。”[41]

那么目前在宇宙常数问题上有些什么看法呢？可以看出有两种流行的观点：一种是极端的观点（如W.泡利所表述的[42] ），认为宇宙项“是

多余的，没有根据的，应当抛弃”；另一种是温和的观点（如W.林德勒所说的[43] ），认为宇宙项“属于场方程，如同一个附加常数属于不定积分一样。”温和观点是可取的，因为Λ项除了在宇宙学研究外并不重要，并且对于人们通常所考虑的那些宇宙模型来说，加上这一项几乎没有增加求解的复杂性。显然，不管怎么说，任何一个严肃的相对论学者都不会同意爱丁顿“令Λ=0会动摇空间的根基”的观点。

Ⅻ

我想以一段趣闻愉快地结束关于爱丁顿工作的这部分介绍。

爱丁顿1924年访问了加州大学伯克利分校物理系。就在那一次，爱丁顿与一位名叫威廉斯（W.H.Williams）的教授共用一间办公室，两人每周去克莱蒙特俱乐部玩两次高尔夫球。在他离去的前一天晚上，教员俱乐部特地为爱丁顿安排了一次晚宴，威廉斯教授应邀发表讲话。正如威廉斯记述的那样：[44]

> ……我曾尽力设法保持严肃，可我还是转到打油诗上。你们知道，爱丁顿是一位《爱丽丝梦游仙境记》迷。这一点加上卡罗尔与爱因斯坦都共有的颠倒，以及我们对待高贵的高尔夫球运动不恭的态度，构成下面这首诗的主题。

爱因斯坦与爱丁顿

球场青青夕阳照，

月光悄悄爬树梢；

球童甜蜜入梦乡，

两个球手还在跑。

流连洞旁不罢手，

紧紧盯住十三篙。

丁顿爵士因斯坦，

你不罢手我追赶；

爱因斯坦98分，

丁顿几分优势占。

不意两人跌洞旁，

低咒几声接着干。

爱因斯坦把话发，

恁多沙子为的啥？

别处无坑此处有，

真是让人心发傻。

这儿如果平平坦，

高呼三声真伟大！

如有七位少年女，

把这球场扫干净，

那可别说我吹牛，

击他十七球准进。

丁顿爵士忙反击：

你的右曲特差劲。

高尔夫球都滚来，

你在哪儿我猜猜；

小球有的高又瘦，

还有一些胖又矮。

有的圆来有的光，

扁的占多真奇怪！

丁顿爵士找话说，

东扯西拉话题多。

时钟量杆立方体，

钟摆摆动为什么？

空间多远才弯曲，

时间之翅可商榷？

苹果落地为什么，

老师常把引力说；

爱因斯坦把手摆，

$G_{\mu\nu}$才是正经果。

丁顿爵士想不通：

还请先生慢慢说。

引力肯定非拉力，

这点已经不稀奇；

空间几乎空如也，

只有时间无空隙。

丁顿原来要相信，

但是心底仍怀疑。
空间四维不能少，
三维之说不可靠；
斜边平方为几何，
传统说法颇堪笑。
平面几何遭折腾，
丁顿如何不悲号？
时间扭曲弯又翘，
光线也不走直道；
苦苦思索冥冥想，
彻然大悟开了窍。
明天你寄信一封，
今日我却早收到。
今日驱车廷巴图，
两倍光速令人舒；
下午四时启的程，
昨日已经返归途。
总算开窍懂了理，
爱因斯坦拇指竖。
水星绕日作公转，

日夜兼程不停站；

奇怪奇怪真奇怪，

想回起点成梦幻！

看来红尘应看破，

世间万事不用干。

过去事情还没完，

未来就已抢先占；

不论女王洋白菜，

都只让人心发寒。

校长还有系主任，

还有啥事值得干？

爱因斯坦又发话，

最短路线像个8；

弯来绕去却最短，

这事一想头发麻。

你想先到走得快，

然而欲速则不达。

复活节日是圣诞，

远近事情一般看；

二加二，大于四，

如此这般不用算。

丁顿爵士忙点头，

云里雾里团团转。

似懂非懂直点头，

谢你耐心说缘由；

我流眼泪你莫怪，

头痛如裂泪自流。

得到头清脑灵时，

接着话头说个够。

XIII

我前面主要讨论了爱丁顿20世纪20年代以前对天体物理学和经典相对论的贡献。1926年，爱丁顿的《恒星的内部结构》出版，那年他44岁。在以后的18年中，除了偶尔涉足早年感兴趣的领域外，爱丁顿专心致力于证明他所选择的宇宙模型的合理性，进而将这一模型作为他“统一量子论和相对论的基本理论”之基础。从实质意义来说，我不敢宣称自己理解了爱丁顿的基本理论。但按爱丁顿自己的话说，他的理论有两个基本前提。据我判断，这两个前提或者不能成立，或者未被接受。可首先，我得尽量客观地解释一下爱丁顿对所述问题在宇宙学方面的见解。

爱丁顿写作《相对论的数学理论》时，有两个宇宙模型[45] 都依赖于非零宇宙常数Λ：静态的、处于流体静力平衡的爱因斯坦宇宙，和静态的却膨胀着的德西特宇宙。两种模型理论上都是可能的，并且满足均匀性和各向同性的假设。德西特宇宙在膨胀着，物体以随距离增大而不断增加的速度后退，但爱因斯坦宇宙不存在这样的膨胀。因为德西特宇宙得到他当时所能得到的一些非常贫乏的观测资料的支持，所以爱丁顿在《相对论的数学理论》中表示，他更赞成德西特宇宙作为天体宇宙的

模型。

爱因斯坦宇宙和德西特宇宙都是静态的，其意思是指所有的度规系数都不依赖于时间。后来人们认识到，尽管德西特宇宙表现出膨胀，其静态特性源于它不包含任何质量密度。因此，爱因斯坦宇宙是当时唯一没有运动的宇宙模型，爱丁顿这样作了总结：

> *爱因斯坦宇宙包含了物质但不包含运动；德西特宇宙包含了运动却不包含物质。*[46]

然而，弗里德曼在1922年发表的论文中证明：爱因斯坦宇宙容许有一个均匀、各向同性的非静态宇宙模型。勒迈特1927年再次独立地发现了同样的结果，并详细讨论了其天文学影响。爱丁顿后来了解了勒迈特的文章，并在自己的著作中广为宣传。既然Λ为可正、可零和可负的参数，于是就可能有多种宇宙模型可供选择。这些模型现在依然是与天文观测进行比较的基础。但从一开始，爱丁顿的兴趣集中在可能模型的一些具体模型上。他喜欢的模型最初时为爱因斯坦宇宙，其质量（M）和半径（R_E）之间有如下关系：

$$\Lambda = \frac{1}{R_E}\text{，和}\frac{GM}{c^2} = \frac{1}{2}\pi R_E \qquad (18)$$

后来由于不稳定性（这一事实已被勒迈特证明），宇宙开始膨胀。本来它也可能收缩；但爱丁顿等人证明最初的凝集更可能导致膨胀而不是收缩。

为什么爱丁顿选择这样一个特殊模型来描述天体宇宙呢？他这样解释：

> *我是一名搜索罪犯的侦探——这位罪犯就是宇宙常数。我知道他的存在，可不知道它的外观；比如我不知道他是矮个还是高个。*[47]

爱丁顿力图弄清他那位“罪犯”的“外观”，采用的方法大致沿着以下

思路。

对于运动在固定电荷电场中的电子来说，其狄拉克方程中与电子质量m_e有关的项为$m_e c^2/e^2$。爱丁顿断定这一项的出现是因为宇宙中所有其他粒子的存在。更准确地说，他认为这是同宇宙中其余电荷“交换能量”经适当平均后的结果；并且他确信，除了一个量级为1的可能因子外，这一项必然是$\sqrt{N}/R_E$，其中N为宇宙中的粒子数，R_E为初始静态爱因斯坦宇宙半径。他据此得到

$$\frac{\sqrt{N}}{R_E}=\frac{m_e c^2}{e^2} \tag{19}$$

可我们也有关系（见方程18）

$$\frac{GNm_p}{c^2}=\frac{1}{2}\pi R_E \tag{20}$$

其中，m_p表示质子质量。根据以上两个关系，我们得到

$$N=\frac{\pi^2}{4}\frac{e^4}{\left(Gm_p m_e\right)^2}=1.28\times10^{79}$$

以及

$$1/\Lambda = R_E = \frac{1}{2}\pi\frac{e^4}{Gm_p m_c^2 c^2} = 1.07\times10^9\text{光年。} \quad (21)$$

爱丁顿认为这些值与观测结果很好地相符，因此他觉得他找到了“罪犯”。从此，他对自己那些观点的可靠性深信不疑。1943年他在都柏林高级研究所的一次演讲中，他说：

> 过去16年里，我从未对自己理论的正确性感到怀疑。[48]

现在应当提出这样的问题：爱丁顿宇宙模型目前地位如何？1944年以来我们认识方面的进展不容置疑地表明，爱丁顿宇宙模型不是一个正确的模型。证据主要来自宇宙3K黑体辐射的发现及氦具有原始起源这一事实。

宇宙中充满了一个均匀、各向同性的辐射场，该辐射场具有3K温度的普朗克分布。这一事实意味着，自宇宙辐射温度为4000K、物质和辐射发生分离到现在，宇宙已膨胀大约1300倍。与此类似，由核聚变产生原始氦说明，宇宙曾处于密度数量级为1000g/cm^3 、温度量级为10^9 K这样的状态。反过来，这些要求又意味着那时宇宙半径必然要小10^9倍。宇宙半径这么大的变化与爱丁顿宇宙模型完全不相容。

1944年6月9日，仅仅爱丁顿去世前5个月，爱丁顿最后一次出现在皇家天文学会会议上，并提交了《星系的退行常数》的论文。这次爱丁顿表示，方程（21）给出的Λ“就是宇宙常数”，他还进一步指出：

> ……宇宙演化的时间尺度肯定比90×10^9 年小，而且在我看来突破这一极限不太可能。[49]

提交这篇论文之后，爱丁顿与麦克维蒂（G.V.Mc Vittie）进行了如下对话：

> 麦克维蒂博士：你的理论似乎完全建立在从爱因斯坦宇宙开始膨胀的模型之上。我们知道，对螺旋星云分布的“深入”观测，为在不同宇宙模型之间做出鉴别提供了一个非常精密的检验标准。如

果正巧观测结果没有选中你所用的特定模型，你的理论又将如何呢？

爱丁顿爵士：我认为，在很长一段时间内，我们得不到足够精确的资料来解决这一问题，所以我觉得没有必要考虑偶发事件！除了直接与观测结果比较以外，我们必须努力寻找其他方法得出结论。[50]

爱丁顿所说的偶发事件终于发生了；麦克维蒂的问题仍有待解决。

ⅩⅣ

我已经对爱丁顿将宇宙中的粒子数与原子常数联系起来的推理思路作了说明。为了进一步发展他的理论，他需要数出爱因斯坦初始静态宇宙有效相空间中体积为h^3的相格数目。在爱丁顿的心目中，要求的数目与依据泡利不相容原理推出简并电子气状态方程时所作的计数，如果不说完全相同的话，也是密切相关的。既然我与这件事有关联，我想谈谈爱丁顿就此问题对自己思想演化过程的详细描述，这段话引自1936年夏爱丁顿在哈佛大学艺术与科学三百周年大会上的报告。

……恒星内1000万度的高温，使得大多数卫星电子摆脱原子的束缚，原子剩余的部分只是一个很小的结构。原子或离子变小了许多，只有当密度增大1000000倍时，它们才会拥挤起来。因为这一原因，恒星密度大大增高时仍然保持着理想气体状态。太阳和其他高密度恒星仍然遵从理想气体的理论，这是很自然的，因为组成它们的物质是理想气体。

因此，没有任何因素能阻止恒星物质浓缩至极高的密度；这本身说明，对于称为白矮星的一类恒星，根据观测计算出来的密度，尽管高得令人难以置信，可也许完全是真实的。

在得出这个结论时，我并非没有疑惑。这些超密恒星最终会怎样，我对此深感不安。恒星似乎陷入进退两难的尴尬境地。最终恒星储存的亚原子能量会耗尽，而这时恒星就要冷却下来。这可能吗？要知道，只有极高温才可能出现惊人高的密度。如果恒星物质冷却下来，按道理它可能会恢复到地球的密度，但这意味着恒星体

积必须膨胀到现在体积的5000倍那么大。膨胀需要能量克服引力做功；恒星似乎没有可用的能量储备。如果恒星继续丧失热量，而又没有足够的能量来使之冷却，那它到底该怎么办！

亚当斯教授及时地证实天狼星伴星具有很高的密度，但这个谜仍然存在。不久，福勒教授出来解围，他在一篇著名的论文中，应用了波动力学中刚得到的一个新成果。巧得很的是，就在天文学中发现超高密度物质时，数学物理学家们正完全独立地将注意力转向同样的课题。我猜测，1924年之前人们没有认真考虑超密物质问题；可正当它在天文学中出现时，它也同时出现在物理学中。福勒给出了证明，刚发现不久的费米—狄拉克统计法把恒星从我一直担心的不幸命运中解救出来。

物理学家们并不满足现状，他们着手对福勒公式加以改进。他们指出，就白矮星的条件而言，电子速度接近光速，福勒忽略了某些相对论效应。称作普通简并公式的福勒公式，逐渐被一个新的公式所取代，这个新的公式称为相对论简并公式。似乎一切都好，直到钱德拉塞卡的某些研究说明了这样一个事实：相对论简并公式将恒星带回到与福勒将它们解救出来之前完全相同的困境中去。小的恒星可以冷却下来，并以某种合理的方式结束一生，变成暗星。但超过某个临界质量（太阳质量的2倍或3倍）的恒星，永远也不会冷却下来，而必须不断辐射和收缩，最后上帝才知道它要变成什么。这并未让钱德拉塞卡担心；他好像喜欢恒星如此变化，他相信实际情况就是如此。可我早在12年前就对这种恒星小丑行为有同样的恶感；至少它的怪异足以引起我的怀疑：一定是所用的物理公式出了毛病。

我对相对论简并公式做了仔细检查，我发现：它是相对论和非相对量子论联合作用的结果。我不认为这样一种结合的后代会是合法婚生的。相对论简并公式——现在所用的公式——实际没有任何根据，正确运用相对论得出的公式应是普通公式——即大家已经抛弃的福勒原始公式。这一点或许令人相当意外。我发觉宣布这样的结果无异于捅马蜂窝，我对此毫不奇怪。物理学家们在我耳旁嗡嗡直叫，可我认为自己还未被咬伤。无论如何，就这次演讲的目的而言，我想我还没有失言。

我冒昧提到这项研究个人方面的情况，因为它表明科学的一些不同分支之间是如何密切结合在一起的。就我而言，当钱德拉塞卡的结论引起我对相对论简并公式的怀疑时，再为此花费时间做深入研究是困难的，因为我正致力于属于不同思想领域的一项长期研究。这项工作已经花了我6年的时间，正接近完成，只剩下一个问题就使之圆满结束，余下问题就是宇宙常数的精确理论计算。但这个问题完全把我难住了。不过我刚得到4个月的闲暇时间，我打算利用这段时间解决这个问题，或者说尽最大努力。可一旦开始考虑简并公式，我就难以自拔。这占去了我全部的时间。时间悄悄逝去，我把宇宙常数问题完全放下了。有一天，我试图根据各种观点检验我得出的简并结论时发现，在一种极限情况下，它变成一个宇宙学问题，这为我已搁置一旁的宇宙常数问题指出一个新的方法——采用这个新方法，宇宙常数问题没有太大困难就得以解决。我现在终于明白，用其他方法将很难解决；如果将4个月时间花在原已计划的直接进攻上，我不可能取得任何进展。

几天前。我在数学部宣读了一篇论文，其中给出了螺旋星云退行速度和宇宙粒子数目的计算。这篇论文源于天文学，可它并不是由于考虑螺旋星云提出来的。它产生于对天狼星伴星和其他白矮星的研究。[51]

让我解释一下争论要点是什么。福勒关于白矮星物质状态的讨论基于完全相同的简并电子气理论，该理论因索末菲的金属电子理论而广为人知。支配这样一种电子气的状态方程为

$$p=\frac{1}{20}\left(\frac{3}{\pi}\right)^{2/3}\frac{h^2}{m_e}n^{5/3} \qquad ㉒$$

其中，p表示压强，n表示每cm^3 的电子数目。可是，在白矮星中心处通常的密度条件下，处于费米阈值的电子具有接近光速的速度。如果按照当时即已普遍且至今仍在使用的方法考虑这种情况，人们就会发现，状态方程并不遵从方程（22）所给的形式，在电子浓度很高的极限情况下，状态方程趋向于

$$p = \frac{1}{8}\left(\frac{3}{\pi}\right)^{1/3} hcn^{4/3}\left(n \to \infty\right) \tag{23}$$

爱丁顿就是认为对状态方程的这一修正是“毫无根据的”。

非相对论形式的状态方程为（22），精确形式的状态方程在低密度极限情况下为（22），高密度极限下为（23），采用以上两种形式的方程会导致如下结果。

根据非相对论状态方程（22），处于平衡状态下的恒星，其半径与质量的立方根成反比。因此，对任意质量来说，有限平衡结构都是可能的。爱丁顿正是认为这一事实特别令人满意。然而，若采用高密度极限下精确形式的状态方程（23），人们发现，如果质量超过以下极限，任何平衡状态都不可能达到：

$$M_{\text{limit}} = 0.197\left(\frac{hc^{3/2}}{G}\right)\frac{1}{\left(\mu_e H\right)^2} = 5.76\mu_e^{-2} \tag{24}$$

其中，μ_e 为电子的平均重量。人们得到完整的质量—半径关系[52]，如图4所示。

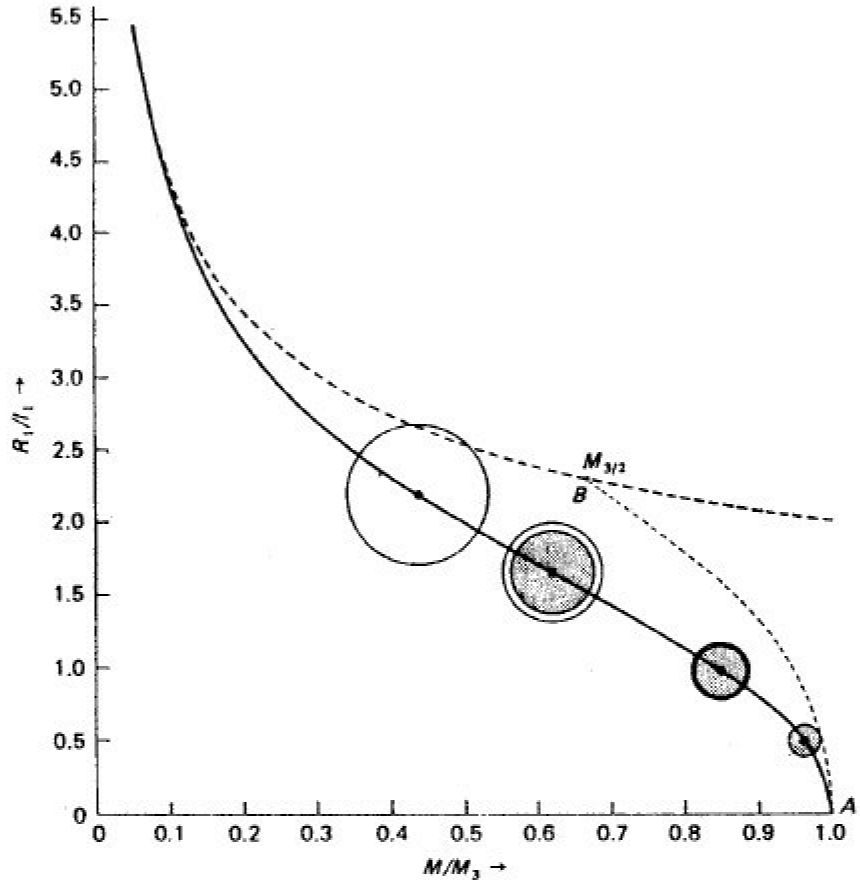


图4实曲线表示完全简并结构的精确（质量—半径）关系。横坐标为质量，以极限质量（用M3表示）为单位，纵坐标为半径，以l1=7.72μ¯ 1×108cm为单位。虚曲线表示由状态方程（11）得出的关系；在这条曲线的B点，结构中心处电子的阈动量P0正好与mec相等。精确曲线上画了一个整圆（未带阴影）的地方，P0（中心处）再次与mec相等。其他圆圈的阴影部分代表了这些结构中电子可被认为是相对论性电子（P0＞mec）的那些区域

质量超过极限（24）时，根本不存在任何有限简并态恒星结构。爱丁顿正是认为这一事实是“恒星小丑行为”。如同他早先所说的那样：

> 钱德拉塞卡利用5年前已被接受的相对论公式表明，质量大于某一极限M的恒星仍然是理想气体，它不会冷却下来。恒星不得不持续不断地辐射和收缩，直到最后，我猜想，其半径变成几千米，此时引力足以抑制辐射，恒星最终能够平静下来。
>
> 钱德拉塞卡博士以前就得到这一结果，可他在最近的一篇文章中反复提到它；与他讨论这个问题时，我觉得没法不得出结论：这一结果简直就是相对论简并公式的一种归谬法。可能会出现各种偶然事件挽救恒星，可我更想得到一种强有力的保护机制。我认为应

该存在一条自然法则，阻止恒星发生这种荒谬的行为！[53]

从这段叙述可以清楚看出，爱丁顿早在1935年就完全认识到，简并结构质量如果存在上限，就必须考虑到导致形成我们所称黑洞的引力坍缩之可能性，可他不愿接受他自己预见中已经得出的结论，却要确信“应该存在一条自然法则阻止恒星发生这种荒谬的行为！”

此外，我们应该问一问，爱丁顿就这一问题做出的判决现在地位如何呢？简单而又直接的回答是：它们未被接受。现在，极限质量的存在已经与恒星演化和某些恒星高密度核心核燃烧的复杂设计、超新星的引力坍缩、几乎具有相同质量的中子星及黑洞的形成一起，不可避免地被编织进现代天文学织锦之中。但我们现在还不是时候，也没有时间谈这些。即使是对最粗心的观测者来说，所有这些都是很容易看出来的。至于我个人，我只想说自己很难理解，作为广义相对论最早最忠实的支持者，爱丁顿为什么会觉得恒星演化的自然过程中会形成黑洞这一结论是那样不可接受。好像一承认黑洞理论，他所建造的宏伟理论大厦中的两个主要支柱就会坍塌。暂时抛开这一事实不谈，从这所大厦本身我们能得到些什么呢？这里有两种看法。

> 爱丁顿的工作，如果正确的话，那是极端重要的，可大多数试图了解其工作的人并没有同意他的结论。他的一些论文，在讨论的某一点上很清楚，接着在关键时刻它们又含糊起来，当推出一个重要结论以后重又清晰起来。所得的结论自然无法通过那些叙述明了的公理和假设合乎逻辑地推演出来。爱丁顿本人对此十分清楚。笔者有一次与他做了一次长时间的讨论后，仍未取得任何进展，于是爱丁顿说：“尽管我无法给出证明。可我确信我的结论是正确的。”（A.H.Wilson）[54]

> 爱丁顿的“统一理论”，除了它的含糊性外，它解释了太多的东西：事实上它对一切都作出了解释，到头来人们发现这种理论什么也没有解释。这种理论充其量只是未完成的工作，正如达·芬奇的科学研究一样，它所包含的洞察力的闪光，可能要等到其中的重要内容通过非常不同的途径和方法被充分发掘之后，才会为后代人所赏识。（J.G.Crowther）[55]

然而，几乎毫无疑问，即使爱丁顿的大厦部分地崩溃，可有些高高

的石柱仍然耸立着。举一个例子就够了。在处理狄拉克方程时，爱丁顿改进一种E数演算，它本质上是包含16个元素的群代数，满足狄拉克矩阵的反变换规则。这一改进本身代表一项了不起的成就，它为现在重新唤起对克利福德代数的兴趣，产生了重要的影响。爱丁顿改进的基本元素如下。

爱丁顿一开始以下列表式定义了5个E数：

$$E_1=\begin{vmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{vmatrix},\quad E_2=\begin{vmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{vmatrix},\quad E_3=\begin{vmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{vmatrix},\qquad (25)$$

$$E_4=\begin{vmatrix} i\sigma_2 & 0 \\ 0 & -i\sigma_2 \end{vmatrix} \text{和}\ E_5=\begin{vmatrix} 0 & i\sigma_2 \\ i\sigma_2 & 0 \end{vmatrix},$$

其中，σ_1，σ_2 和σ_3 为泡利2×2矩阵。这些E数满足交换律，

$$E_\mu E_\nu+E_\nu E_\mu=-2\delta_{\mu\nu}\ (\mu,\ \nu=1,\ \cdots,\ 5) \qquad (26)$$

和

$$E_1E_2E_3E_4=iE_5 \qquad (27)$$

爱丁顿代数的16个元素为：

$$i,\ E_\mu \text{和} E_\mu E_\nu\ (\mu,\ \nu=1,\ \cdots 5;\ \mu\neq\nu) \qquad (28)$$

爱丁顿进一步将E代数的平方定义为16×16的复合短阵。这个“双E框架”与9维基本空间中三个克利福德代数之一相对应。由于近来关于超

对称规范场的工作是基于8维和9维克利福德代数的，从这一点上说，爱丁顿远远地走在时代的前面。

爱丁顿认为找到等幂E数（即满足E^2 =E）的代数元素很重要。它们是量子电动力学中动量和自旋的投影算子。最后，E代数在实数范围内是五维的，这一认识导致他（在粒子物理学中第一次）引入了“手征”（chirality）的概念，这起源于方程（27）右边选择+i或-i。还应指出的是，爱丁顿最先认识到4×4实矩阵代数及其重要性。这一代数后来由马约拉纳发现；物理学文献中常常称之为“马约拉纳旋量”（Majorana spinors）。[56]

ＸＶ

我想简要回顾一下爱丁顿的著作，总结他关于科学研究态度的变化。如果我们将他的这些态度与他研究的方向做一对照，或许能够对其力量与弱点之根源有所了解。

1920年，在他对英国学术协会所做的关于“恒星的内部结构”演讲中（前面我已引述过），爱丁顿以相当的篇幅讨论了科学研究中思辨和理想化模型的作用。他是这样说的：

> ……什么是我们可以用来检验科学理论合理发展、抛弃无用思辨的试金石呢？我们都知道，一些理论被科学思想作为无用的思辨而本能地抛弃；但要指出它们的缺陷，或者提供一条能够判断我们自己什么时候犯错误的规则来，却非常困难。人们常常以为思辨与假说是一回事，但更多的时候它们是对立的。只有我们让自己的思想偏离古老但有时并不可靠的假说时，我们才被认为在进行思辨。假说限制思辨。而且，对思辨的怀疑可以弥补思维不严密这一缺陷——一些混乱的思想在我们心中抛了锚，会影响我们的视野；另一方面，思辨性太强的话，就没法让它们接受彻底的科学检查，驱除其中的邪魔。
>
> 如果我们不满足于经验事实的单调积累，如果我们想进行一些演绎或推广，如果我们要寻找理论来指导我们，一定程度的思辨就无法避免。有些人喜欢采取似乎非常直接的解释并立即将它作为一个假设，另一些人则设法探求出与事实不是完全矛盾的最广泛的可

能性，并进行分类。两种选择都存在着危险：前一种选择也许太狭隘，将结果导入死胡同；后一种可能太广泛，作为一种指导思想没什么用，并与经验知识长期脱节。当后一种情况发生时，一定会得出这样的结论：知识还不适合于进行理论处理，思辨的时机还不成熟。思辨性理论和观测性研究可互相配合、互相促进的时候，正是可能性——或至少是概率——可通过实验缩小，理论能够指出检验方法的时候，这些检验可将遗留的错误路线一条一条地堵塞掉。

理论物理学家处境极其困难。他可以设计一个理想物质模型，该模型的行为具有特定的属性，数学上遵从严密的定律。他的工作到此为止是无懈可击的，与二项式定理相比，它并不更思辨：但当他断言他的玩意儿如何如何了不起，当他指出他的模型恰好反映了自然界的某种行为时，他就真正地陷入了思辨。客观实体果真像理想模型那样吗？会不会有另外的未知条件介入呢？他没有把握，但不能不进行对比；因为只有不断地观察自然才能帮助他选择目标。他经常会犯下的一个共同错误就是，使用那些更有经验的观测者对之摇头的数据资料作依据，其实它们太不可靠，不能作为广泛比较的依据。然而即使在这种情况下，理论仍然可以指出什么类型的数据特别重要，从而对观测给予帮助。

我认为研究若能得到正确观点的指导，那些空洞的思辨就可以避免。如果利用严格的数学得出理想模型的特性，所有基本的假设都被清晰地理解，这时就有可能说什么什么特性和定律精确地导出什么什么结果。如果存在某些未被考虑的其他因素，它们会在与自然的对比中暴露出来。如果模型未能与观测达成完全一致，没有必要对此失望；模型已经完成其使命，因为它暴露出客观现象的特点是什么，这些现象需要新的条件才能解释。与观测达到初步的一致是必要的，否则模型就毫无希望；并不是说在目前范围内它必然是错的，而是说它明显地将次要的特征突出来。如果我们一直是扯着死结的错误一端，死结就永远解不开，这时只有通过别的途径方能解开它。但当与观测之间大略一致后，死结开始松动，我们常常得为下一个困难做准备。我认为，自己的理论经受住一次更严格的观测检验后，应用数学家不应为此满足，而应感到失望——“又一次受挫！这次我本希望发现有不一致的地方，它能指出我的模型在什么地方还可以得到改进。”或许这是无法做到的理想化建议；但我承认我从未强烈感受过这种失望。

> 我们的自然模型不应像一座建筑——一座供大众欣赏的漂亮的建筑物，总有一天，抽去它的一块基石后，大厦会倒塌下来。它应该类似于一台部件可以移动的机器。我们不必固定每一杠杆的位置，它应该根据最新观测结果不断进行调整。理论家的目标就是要弄清由杠杆带动的轮系——也就是作为机器灵魂的各部件的联结系统。[57]

上面这段叙述。几乎任何严肃、老练的天体物理工作者都不会反对。爱丁顿以谦虚的评价结束了论“恒星能量的来源”的演讲：

> 我本想把话题一步一步地引向某个伟大的高潮，从而结束这些演讲。可是，或许与科学进步的真实情况更相符的是，好的结尾倒是应该虎头蛇尾，其中还应闪耀新知识的曙光。我无意为结论的无说服力道歉，因为它不是结论。但愿我确信它还只是开始！[58]

这与我所了解的科学探索方法完全一致：但他两年后说的话反映出他的态度有了变化：

> 在科学中，我们有时对一个问题的解十分珍爱，它不能被证实，但我们却相信它一定正确；我们受到某种对事物合理性的先天意识的影响。”[59]

不久，爱丁顿对于他关于宇宙常数、宇宙模型、相对论简并、黑洞的形成以及“统一量子论和相对论”方法的观点，变得过于自信。从我以上对其著作和演讲的各种引述中，可以很清楚地看出这一点，将他1926年对他自己在恒星内部结构方面的工作所作的谦虚评价，与10年后他对我作的过于自信的评论相比较，爱丁顿态度上的这种急剧变化就显著地表现出来了。

> ……你从恒星的观点来观察它；而我从自然的观点观察它。

很清楚，爱丁顿这时对自己的观点已经过分地自信了。

尽管爱丁顿表达了对其基本理论正确性的信心，可是，同时代人对他的工作的忽视，一定使他深深感到沮丧。在他1944年写给丁戈尔（Dingle）的一封哀怨的信中，表现出他的这种遭受挫折的心情

我一直试图弄清为什么人们觉得我的理论含糊不清。但我要指出，尽管爱因斯坦的理论被认为是含糊的，但是成百上千的人认为有必要把这种理论解释清楚。我不相信我的理论比狄拉克的理论更深奥晦涩。可对于爱因斯坦和狄拉克，人们却认为克服困难去弄懂这种深奥晦涩是值得的。我相信，当人们认识到必须理解我，并且“解释爱丁顿”成为时尚时，他们也会理解我的。[60]

以下这段话读过之后令人难以忘怀：

在他最后的日子里，由于长时间地想入非非，他脸色如死人般苍白，显得痛苦不堪。[61]

1920年，在英国学术协会演讲中，爱丁顿讲述了蒂达洛斯（Daedalus）和伊卡洛斯（Icarus）的故事，或许当时他已预见到他未来科学探索的方向：

古代有两位飞行员，他们给自己装上翅膀。蒂达洛斯在不太高的空中安全地飞越了大海，着陆的时候理所当然地受到赞誉。年轻的伊卡洛斯迎着太阳高飞，最后捆绑翅膀的蜡熔化了，飞行也就因此彻底失败。在衡量他们的成就时，也许要为伊卡洛斯说几句话。第一流的权威们告诉我们说他仅仅是在“玩特技飞行表演”，可我更愿意这样看：是他明确地暴露出他所处时代的飞行器在结构上存在着缺陷。所以，在科学中，谨慎的蒂达洛斯也会将自己的理论应用到他确信非常安全的地方，可它们的潜在弱点不会被他的过分小心揭露出来。伊卡洛斯会将其理论拉至强度极限，直到脆弱的接合点裂开。仅仅是做一次壮观的特技表演吗？也许有几分道理；他也是一个普通的人嘛。但是，虽然他命中注定到不了太阳，无法彻底解开飞机的构造之谜，但我们可以指望从他的失败中得到一些启发，去建造一个更好的飞行器。[62]

因此，在今天，我们怀着极其崇敬的心情，纪念一位曾朝向太阳勇敢高飞的伟大灵魂。

注释

[1]Astrophys. J 101，133（1945）.

[2]A. Wibert Douglas.TheLifeofArthurStanleyEddington（London：Thomas Nelson&Sons.1957）.p.103.

[3]A. S.Eddington, Stars and Atoms（Oxford：Clarendon Press.1927）.p.24；and J.G.Crowther, BritishScientistsoftheTwentieth Century chup.4（London：Routledge&Kegan Paul，1952），p.177.

[4]A. S.Eddington, New Pathways in Science（Cambridge：Cambridge University Press，1935）.p.207.

[5]New Pathways in Science, p. 170.

[6]A. S.Eddington, ScienceandtheUnseenWorld（London：Allen&Unwin.1929），p.33.

[7]Science and the Unseen World, pp. 54—56.

[8]A. S.Eddington，“Forty Years of Astronomy.”in Backroundto Modern Science, ed.J.Needham and W.Pagel（Cambridge：Cambridge University Press.1938）.pp 120—121.

[9]A. S.Eddington, Internal Constitution of the Stars（Cambridge：Cambridge Universisty Press），pp.15—16，245.

[10]Observatory 43. 353—355（1920）.

[11]Nature. 1 May 1926（supplement），no 2948，P.30.

[12]Astrophys J. 101，134（1945）.

[13]Observatory 52. 349（1929）.

[14]Observatory 49. 250.335（1926）.

[15]Observatory 40，224—226（1917）.

[16]Notes&Records Roy. Soc.London 30，249（1976）.

[17]A. S.Eddington, Space Time and Gravitation（Cambridge：Cambridge University Press，1920）.pp.113—114.

[18]“Forty Years of Astronomy，”pp. 140—142.

[19]Mon. Not.Roy.Astr Scc.85，672（1924—25）.

[20]Observatory 42. 989—398（1919）.

[21]Observatory 55，5（1932）.

[22]Observatory 46，142（1923）.

[23]New Pathways in Science, p. 121.

[24]Math. Gazette 19，256—257（1935）.

[25]More generally, if A, B，C, and D speak the truth with probabilities a, b，c, and d, respectively, then the solution of the same problem, as L S. Left-which showed in considerable detail（Math.Gazette 20，309—10〔1936〕），is：

$$\frac{d-acd(1-b)}{1-a(1-b)+\mathrm{a}(1-b)\ (c-2cd+d)}$$

[26]A. S.Eddington, Mathematical Theory of Relativity（Cambridge：Cambridge University Press，1923），p.10.

[27]Observatory 46，193（1923）

[28]Nature 113. 192（1924）.

[29]Proe. Roy.Soc.（Land）A102，268（1922）.

[30]Proc. Roy.Soc（Lond）A166.465（1938）.

[31]The Life of Arthur Stanley Eddington. pp.189—192.

[32]Mathematical Theory of Relativity, p. 153.

[33]Ibid, p. 154.

[34]Ibid, p. 155

[35]A. S.Eddington, TheExpandingUniverse（Cambridge：Cambridge University Press，1933）.p.35.

[36]New Pathways in Science, p. 315.

[37]The Expanding Universe. pp.147—148.

[38]Einstein：A Centenary Volume, ed. A.P.French（Cambridge Mass：Harvard University Press，1979），p.274.

[39]The Life of Arthur Stanley Eddington. p.57.

[40]Proc. Nat.Acad.Sci.18，213（1932）

[41]“Forty Years of Astronomy.”p. 128.

[42]W. Pauli, Theory of Relativity traps.G Field（London：Pergamon Press 1958），p.220

[43]W Rindler, Essential Relativity（Berlin：Springer-Verlagt，1977）. p.226.

[44]RecordsofR. A.S.Club 1925—1953，ed G.J.Whitrow.pp.xxiv—xxvii

[45]I am excluding the nonstatic solutions discovered by Freidmann in 1922 since they came to be known generally only several years later.

[46]Proc. Phys.Soc.44.6（1932）

[47]The Expanding Universe. p.87.

[48]Dublin Inst. Adv.Studies A2，l（1943）

[49]Observatory 65. 211（1944）.

[50]Ibid. p.212.

[51]Ann. Rep.Smithsonian Institution（Washington, D.C：U.S.Government Printing Office，1938）.pp.137—139.

[52]Mon. Not.R.Astr.Soc.95，207（1935）.

[53]Observatory 58，37（1935）

[54]Cambridge Review 66，171（1945）.

[55]British Scientists, p. 195.

[56]I am greatly indebted to Dr. N.Salingaros for drawing my attention to the significance and the importance of Eddington's work on his E-numbers.

[57]Observatory 43，356—357（1920）.

[58]Nature，1 May 1926（supplement），no 2948，p. 32.

[59]A. S.Eddington, The Nature of the Physical World Cambridge：Cambridge University Press，1928），p.337.

[60]British Scientists, p. 194.

[61]Ibid, p. 143.

[62]Observatory 43，357—358（1920）.

爱丁顿的其他著作

Stellar Movements and the Structure of the Universe（London：

Macmillam&Co. 1914）.

Report on the Relativity Theory of Gravitation（London：Physical Society，1915）.

Relativity Theory of Protons and Electrons（Cambridge：Cambridge University Press，1936）.

Fundamental Theory（Cambridge：Cambridge University Press，1946）.

第7章　K.史瓦西讲座广义相对论的美学基础（1986）

Ⅰ

卡尔·史瓦西无疑是本世纪最杰出的物理学家之一。他的贡献之多令人惊讶：物理学、天文学、天体物理学的广阔领域到处都留下了他的足迹。

在物理学领域，从电动力学、几何光学到当时最新发展起来的玻尔和索末菲的原子理论等方面他均有贡献。在电动力学方面，他得出了洛伦兹电子方程变换的基础；在几何光学方面，他发展了光学仪器的象差理论（后来玻恩认为它有“空前的明晰性和严格性”），并表述为基本原理，这一原理成为施密特（Schmidt）望远镜光学的理论基础；在玻尔—索末菲理论方面，在他生前发表的最后一篇文章中，得出了双原子分子转动—振动光谱的斯塔克（Stark）效应和德朗德尔（Deslander）项（在这篇文章中，他第一个引入了作用量和角变量的概念）。

在天文学和天体物理学领域，史瓦西的贡献太多太广，这里我只能提及以他的名字命名的一些发现。照相光度学的史瓦西指数；辐射传递理论中的史瓦西—米尔恩积分方程；对流不稳定性开始的史瓦西判据；星球速度的史瓦西椭球分布；当然，还有描述质量呈球形分布的球外和静止黑洞外时空的爱因斯坦方程的史瓦西解。这些成果都是在短短的20年间取得的！

我的报告题目可能会使一些人感到困惑：它与这次系列讲座前面讲的有些不同，它不针对某一具体的课题；我担心它在天文学方面没有多大的意义。然而我的报告将从史瓦西发表的文章中探讨他对于科学问题的态度和研究方法；而且，还将直接谈到他的广义相对论方程的解。

Ⅱ

这里，我将从史瓦西研究活动中举三个例子来说明他是如何研究科

学问题的。

第一个例子是他关于星流（star-streaming）方面的研究。星流是卡普坦发现的。并立即由爱丁顿用他和卡普坦的二星流假说得到了恰当的解释。史瓦西对这个假说的看法是：

> 在二星流中星体固有运动的大小是相等的，还可以假定它们离开太阳的平均距离也相等。所以，二星流中的星体，他们在彼此相对运动的过程中，必定有共同的涨落，这种涨落是如何发生的，现在还搞不清楚。
>
> 所以，我相信爱丁顿自己做出的相同的观察资料，应该重新整理，以便得出一个关于星体运动更统一的假说。

以此为依据，史瓦西得出了恒星特有的速度的椭球分布公式。这个公式后来成为所有有关恒星运动和星系动力学理论的基础。然而，更使我注意的是他的这一观点：对自然的描述必须自然，不能是特设的。

第二个例子，我将从他更早发表的文章中选取。在1900年的一次学术会议上，史瓦西讲了这样一个问题：天文学的三维空间几何是否可能是非欧几里得几何。他说：

> 你一定知道，19世纪有人在欧几里得几何以外提出了非欧几里得几何，其主要实例就是所谓的球面和赝球面空间。我们如果知道在可能具有有限曲率半径的球面或赝球面几何中世界是什么样子，我们会感到惊异。如果有这种可能，你会感到自己处在几何学的仙境里；而且如此美妙的仙境会不会变为现实，我们也无法知道。

我们不能不对史瓦西的科学想象，以及他在广义相对论创立之前15年就论述这样的问题而感到惊奇，但对史瓦西来说，这是最简单不过的想象。他实际用他那个时代得到的天文数据估算了三维空间曲率半径的极限，并得出如果空间是双曲形的，其曲率半径不可能小于64光年；如果空间是球形的，其曲率半径至少为1600光年。

我们大可不必为史瓦西具体估算的数字而争论，这儿更有意义的是，史瓦西把他的想象力发挥得淋漓尽致，并使他进入了一个仙境般的世界！

第三个例子是史瓦西关于爱因斯坦真空方程解的发现，它是质量以中心球对称分布时的外部解。毫无疑问，这是广义相对论创立后的一个最重大的发现。

1916年1月13日，爱因斯坦把史瓦西导出真空方程解的论文递交给柏林科学院。仅仅在两个月以前，爱因斯坦才在一篇短的通讯中公布了他的广义相对论基本方程；而全面详细推导的论文，在6个月以后才发表（在这篇论文中，爱因斯坦从理论上推出了水星在近日点的进动速率和一束光掠过太阳边缘时偏折的大小）。1916年1月9日，在致史瓦西的信中，爱因斯坦写道：

> *我怀着极大的兴趣拜读了你的论文，我没有料到有人竟能用如此简洁的方法得出这个问题的精确解。你对这一问题的分析处理实在是妙极了！*

在当时的环境下，史瓦西得出现在这一著名的解是非常了不起的。1915年春夏期间，史瓦西在德军东部前线服役，任很不起眼的技术参谋。在服役期间，史瓦西得了致命的天疱疮，并于1916年5月11日与世长辞。就在他生病期间，他写了两篇有关广义相对论的文章。第二篇中解决了均匀质量的平衡问题，并得出如果物体的半径小于9/8史瓦西半径，流体静平衡是不可能的。这一半径就是我上面提到的玻尔—索末菲理论中的基本量$2GM/c^2$ 。

关于史瓦西弥留之际的病情，爱丁顿在催人泪下的讣告中写道：

> *他的结局是悲惨的，他以极大的勇气和毅力长期忍受着战场上患下的可怕的病痛带给他的折磨。*

这里顺便插一句，理查德·柯朗（Richard Courant）在30年代末曾告诉我。在战争期间他作为参谋本部的成员正从东部前线随部队撤下来，正好遇到史瓦西正向东部前线行进，他感到奇怪，像史瓦西这样有名望的人竟然被派往在参谋本部看来十分危险的前线去。

还是回到史瓦西的论文以及他力图精确求解那个爱因斯坦问题的缘由上来。本来爱因斯坦以前已用近似方法得出了一个解，对此，史瓦西是这样认为的：

用一种简便的形式得出精确解总是一种满足。在现在情况下，确立一个唯一的解并消除所有爱因斯坦关于处理这问题方法的疑虑就更加重要了。因为，正如下面所显示的，在这一问题上，确定一个近似过程的有效性是十分困难的。

在早期爱因斯坦给史瓦西的信中，曾争论过这一问题，爱因斯坦认为，解这个方程近似方法的有效性是不用怀疑的。而史瓦西意识到，精确求解这一问题对新建的理论是很重要的，大胆去求精确解也是十分有意义的。我说“大胆”是因为精确解决这一问题颇费力气，因为人们很快就发现，爱因斯坦广义相对论中有关这个问题的困难，特别是它的精确解更令人棘手（几十年后，才发现它有害于理论）。

过一会儿我再谈这个精确解在理解广义相对论中扮演着什么样的角色。现在我必须转到我这个讲演的主要目的上去。

III

人们常常把广义相对论描绘成一个非常美的理论，甚至把它比喻为一件艺术品（例如玻恩和卢瑟福）。狄拉克（Dirac）毫不过分地说：

是什么使物理学家们这样乐于接受这一理论的呢？那就是它伟大的数学美，尽管它违背了简单性原则。（1939）

爱因斯坦的引力理论本身具有优美的特征。（1978）

以上这些评论和类似的一些对广义相对论的刻画引出了如下问题：

广义相对论的美学基础是什么？更重要的是，既然对美的敏感能导致对这一理论的表述和解有更深刻的理解，那么这种对美的敏感应达到什么样的程度？

回答这些问题，需要我们进行深入研究。首先必须意识到广义相对论现有的特殊地位，即在观察和实验中它被证实的情况；其次，尽管经验的支持尚不太充分。但对它却有一种令人鼓舞的自信，这是为什么？

在过去的20年时间里，许多值得称赞的努力花在证明相对论的预测，即证明牛顿理论和自然现象中出现的最小一级偏差上面，这些努力

也是成功的。关于在不同引力位置上时间具有不同的变化率；关于光线经过引力场发生弯曲及随后的时间延迟；关于开普勒椭圆轨道的进动；还有，由于引力辐射的发散，双星的偏心圆轨道周期逐渐加长。广义相对论的所有这些预言都在观察和实验误差允许的范围内得到了证实，但与牛顿理论所预言的仅相差百万分之几，只不过是爱因斯坦场方程中的后牛顿展开中的三四个参量而已。但是，至今在强引力场范围内，广义相对论的预言还没有得到任何证实，似乎在可预见的将来一段时间内也没有。

一个理论把原先的理论扩大了，而且在原来范围里，它能如同牛顿理论那样有效。那么对于这一理论的证实是归因于与该理论主要方面有关的预言，还是归因于被替代理论与新理论有微小的一级偏离呢？对于这一点我们应不应该搞清楚？例如，难道狄拉克电子理论今日的地位仅在于它成功地解释了1916年帕邢关于电离氦线分裂引起的精细结构的测量结果吗？事实上，狄拉克这一令人鼓舞的理论真正证实，是根据该理论发现了宇宙射线簇中正负电子对。类似地，对于麦克斯韦电磁场方程，如果没有赫兹所做的电磁波的传播速度恰好就是光速的实验，没有庞加莱证明电磁场方程的不变性满足洛伦兹变换，我们会信赖它吗？同样，广义相对论的真正证实只有这一理论特征性预言得到验证，才会很快实现。黑洞作为巨星自然演化最终平衡态的出现，从任何真正的意义上看，不是广义相对论预言的证实。光线不能逃脱足够强的引力场的概念仅是一个推论，而不是基于理论的任何精确预言；它仅仅依赖于光线受到引力影响这一经验事实。另一方面，既然广义相对论对黑洞周围的时空进行了精确的描述。也仅当黑洞周围的时空度规得到验证，才能认为真正地“确立了”广义相对论。著名的克尔双参量解是天体宇宙中出现的静态黑洞的唯一解，但对克尔时空度规（或它的另一些方面）的验证，在近期内还不能期待得到。

也许，这里我得指出，如何才能最后验证一个旋转的克尔黑洞的周围时空。人们可以想象，克尔黑洞在赤道上有一个自由电子吸积盘，从中发射出来的偏振光在通过黑洞强大的引力场后，显示出极不均匀的分布，我们能把它描绘下来。慷慨的自然界会不会不吝啬地给出一个能描绘这一图景的清晰的例子呢？这恐怕是我唯一的一次谈及了天体观测。

Ⅳ

正如我所讲的，到目前为止，在观测中，有关广义相对论的精确特征还没有任何验证，在可以预料的将来一段时间里也不会有。那么，为什么我们对这一理论充满信心和信任它呢？有些人曾经说这起因于该“理论对自然描述的数学美”。这种回答是不够令人满意的，还应该说得更详细一些才好。

华兹华斯说：

值得永远信赖的思想，必需建立在坚实的自然根基上。

广义相对论没有坚实的根基，我们的信任建立在什么之上呢？我们的信任是建立在这个理论的内部自洽上，它与我们信赖的物理学要求一致；更重要的是，在构造这个理论时还没有考虑到的一些物理学领域，后来证实与它没有矛盾。我举一些例子来说明这点。

物理学定律的因果性要求，在给出全部类空间三个面的初始值后，在以指向未来的零射线（它们从空间部分的边界发出）为边界的时空区里，未来是唯一确定的。更正规地说，任何物理基本方程，都必须允许一个初始值的描述，这一初始值可唯一地确定以后依赖于初始值的那部分空间的全部情况。场方程也同样有这一性质，为了证实这一点的工作进行得不顺利，直到1944年利希勒诺维茨（Lichnetowilz）才对此做出了证明。

第二个例子，考察在物理学中处于中心地位的能量概念。在物理学中，人们习惯定义一个总体上守恒的局域能。事实上，一个孤立的运动物体将发射引力波，引力波具有能量，这就意味着，在广义相对论中我们不可能期望有一个能量的局域定义。但我们可以指望，如果时空是渐近平直的（从某种明确的意义上说），我们就能从总体上对整个空间（延伸至无穷）定义一个总能量。1962年，邦迪已表明，如果时空在零无限条件下接近平直（即当我们沿零射线走向无穷远），那么，我们就能定义一个质量函数——邦迪质量，它是时间的递减函数；进而，质量函数的递减率精确地等于能量以引力波的形式辐射到无穷远处的辐射率。但是直到最近几年才证明，邦迪质量总是正值，证明需要能量—动量张量T_{ij}满足某些“能量条件”。对理想流体，

$$T_{ij} = (\varepsilon + P) u_i u_j - P g_{ij}$$

这一条件等价于

$$\varepsilon \geqslant | P |$$

上述两例涉及广义相对论的深层结构，说明它内部的自洽性，这种自洽绝不是一目了然的和不证自明的。

广义相对论的一个更加显著的特征就是，只要不超越理论的适用范围，它并不违背物理学其他分支的定律，像热力学或量子力学（我过一会儿讲“广义相对论的适用范围”是什么意思）。

第一个例子考察由狄拉克方程描述的在克尔黑洞时空中的电子波行为。我们知道，从黑洞转动减慢的过程中，我们能得出它的转动能量。更确切地说，如果有一波，其时间t和方位角φ由下式决定：

$$e^{i(\sigma t + m\varphi)} (m = 0, \pm 1, \pm 2, \cdots)$$

频率σ（σ＞0）小于临界值σ_s，

$$\sigma_s = -am / 2M_r + (m = 0, -1, -2, \cdots)$$

这里a和M是克尔参量，r_+ 是视界半径，这样就有超辐射现象（即入射波的反射系数将大于1）出现，这是霍金（Hawking）定理的必然结

果。只要外部辐射源的能量—动量张量与能量正定特征相一致，一个外辐射源与黑洞的每一相互作用，总是导致视界表面积的增加。然而，通过理论的明确数学运算，人们发现，从克尔黑洞反射的狄拉克波，不会出现超辐射现象，显然与霍金定理相矛盾。但是，我们不久就了解到，狄拉克在量子理论中提出的波，其能量—动量张量不满足正能条件。如果按标准的运算可以预言超发光，我们就会得出广义相对论前提条件与量子理论前提条件相矛看，但这种矛盾并没有出现。

我的第二个例子是霍金在1975年证明的，如果根据量子理论把时空曲率看成是电子（或光子）的经典散射势，那么我们应该能从黑洞视界观察到在某一温度下呈费米（或普朗克）分布的电子（或光子）的发射，把这一由视界恒定表面引力所决定的温度和因发射粒子引起的能量减小率联系起来，在形式上得出了熵。当人们对此追根求源时，发现这个熵的概念和热力学统计物理学中熵的概念是完全一致的。热力学和统计物理学并没有期望从广义相对论中得出熵，然而，从这个理论得出的结果并不违背热力学和统计物理学定律。

进一步的证明显示了不仅广义相对论内部自洽，而且它和整个物理学原来适用的范围相一致，这些足以使人们对广义相对论坚信不移了。

V

广义相对论还有一个特征，这与它的美学基础有关。

每一个物理理论的有效性受到它内在条件的限制。例如，牛顿的引力理论就受到物体的速度和光速相比应是很小的限制；经典力学和经典电动力学也受到其相关的作用量要比普朗克的作用量子h大的限制。同样地，我们可以预料广义相对论要受到其时间和距离间隔必须大于普朗克尺度$(\hbar G/c^5)^{1/2}$ 。（约为5.4×10^{-44} 秒）和$(\hbar G/c^3)^{1/2}$ （约为1.6×10^{-33} 厘米）的限制$\left(\hbar=\frac{h}{2\pi}\right)$ 。

任何打破了旧理论的束缚并取而代之的新的物理理论，要特别重视新理论的精确描述是它得以验证的基础。在牛顿的引力场理论中，两体问题的解就是一个范例，它的精确解从定量上解释了开普勒定律。同样地，玻尔的单电子系统理论可以精确地推导巴尔末公式，同时从氢原子的巴尔末系和电离氦的皮克林系的偏离，精确地确定了电子和核的质量

之比。

我们现在要问，广义相对论的新的基本特征是什么？在什么情况下能够清楚明了地揭示这些特征呢？

广义相对论的基本特征是将时间和空间揉为一体这一准确的想法。根据权威的流行说法，这一想法推广了隐含在狭义相对论中的思想。我们要问，是否有这种情形，即新理论中的概念是否已被纯理论证实了呢？时空提供了一个很好的说明。广义相对论极其漂亮地完全解决了黑洞周围的时空问题。黑洞周围的时空非常独特；它非常简单，并且只含两个参量：黑洞的质量和它的角动量。所有已知粒子在这种时空中的行为得到了精确的预言。至今为止，没有一个理论能够提出一个具有如此特征的问题，并能解决得如此简单。广义相对论的这一特征我认为是最优美的几个方面之一，并导致我想进一步探究这一理论的美学基础。

考查一个物理理论并阐明其美学需要的根源，无疑会遇到许多困难。像许多与美有关的讨论一样，它和个人的爱好和气质有关；要使这种讨论具有客观性是很困难的，如果不说不可能的话。但是，对我来说这个问题很有意义。作为一个与广义相对论打了20多年交道的，而且今后还得继续与它打交道的人来说，我可以问自己：这一理论的什么方面对我来说有美感？理论的美学成分是怎样影响并支配问题的叙述和解决的（这些问题使我们对广义相对论的物理和数学内容有更深的理解）。

我已提到黑洞理论。有一个很值得注意的事实：对孤立静止的黑洞，广义相对论得到的唯一解只有两个参量。我在别的场合曾说过：

> 黑洞是一些质量介于几个太阳到几百万个太阳质量之间的宏观物体。在这样的质量范围里，可以把它们看作是孤立的和静止的：它们中的每一个单独的黑洞也因此可以由克尔解精确描绘。这是唯一的一个我们可以精确地描绘一个宏观物体的例子。我们生活中所看到的受各种力支配的宏观物体，可用许多物理理论推出多种近似解。比较起来，黑洞构成的唯一要素就是时空的基本概念。这样，从定义上看，它是宇宙中最完美的宏观物体，既然广义相对论关于黑洞的解仅有两个参量，且解是唯一的，它理所当然是最简单的天体。

但这并不是全部。与所有先前预料的相反，与电磁场、重力场和狄

拉克电子波的传播和散射有关的数学物理标准方程，似及粒子极化光子的测地线方程，所有这些方程都能够被分离和精确地解决。这些方程的分离，导致了人们对有一个多世纪历史的双变量偏微分方程能够分离和解决这一事实，进行了重新检查，一个有意义的数学理论因此而产生了。作为范例，我可以提及在克尔几何中电子的狄拉克旋量方程的分离。作为一个必然结果，它导致在狭义相对论中闵可夫斯基几何球坐标系的狄拉克旋量方程的分离。这种分离在以前被认为是不可能的。

Ⅵ

现在我要转向—个我自己想解决的最困难的问题，也就是我们如何感受广义相对论在数学方面的美，而这种美使得具有物理意义的问题得以阐述和解决。为了不落肤浅俗套，为了更加精确，恐怕我要用到比以前更多的专业术语。

相对论的数学理论主要有两大方面，它们是近些年来发展起来的，即黑洞的数学理论和撞击波（colliding waves）的数学理论。我们已清楚地知道：黑洞是天体演化最后阶段大质量的星体引力坍缩的结果。但是，关于广义相对论中波之间的撞击或散射，却需要加以解释。

在广义相对论中可以构造限制在两平行平面之间单位面积具有确定能量的平面引力波，因此，在极限情形下，我们能够建立一个用δ函数能线图表示的脉冲引力波。附带插一句，在电动力学中是不能建立这样的脉冲波的。能量的δ函数形式意味着场变量δ数有平方根的存在。δ函数的平方根是无法用物理描述的。

1971年克罕（Khan）和彭罗斯考虑两列平行极化的脉冲引力波的碰撞。他们证明，相互撞击的结果来自时空奇点的发展，但这种奇点不像我们已经熟悉的黑洞内部的奇点。这种现象并没有显示理论的任何线性化的形式，通过聚焦撞击波产生的时空奇点与波幅毫无关系。显然，根据上面所说，除了一个精确解之外，是无法揭示新现象的。不管怎样，这个例子中奇点的产生使彭罗斯认识到在广义相对论中，有一个新的领域需要开发。然而，在人们认识到黑洞的数学理论在结构上与撞击波的数学理论有紧密关系之前，这一新领域没有实质性的进展。令人惊奇的事实本身是：人们几乎不会想到处理两个完全不同领域的物理问题所用的数学理论是那么接近。的确，人们在黑洞的数学理论框架上发展起来

的撞击波的数学理论中，发现理论出现了各种各样新的物理意义——这种变化是完全不能料到的。

Ⅶ

要想描述上面我已提及的理论是怎样发展起来的，如果不熟悉广义相对论的语言那是不可能的，这正像我们不熟悉音乐符号就无法分析交响乐一样。

我们关心的是描述静止、轴对称的黑洞时空和描述平面波的碰撞和散射的时空。对于前者，度规系数与时间t和绕轴转动的方位角φ无关，它们只与剩下的两个空间坐标（径向坐标r和极向角θ）有关；对于后者，度规系数与两个类空间坐标x^1 和x^2 （变化范围都从+∞到-∞）无关，它们只与时间t和剩下的空间坐标x^3 有关，x^3 垂直于（x^1 ，x^2 ）平面。

可以证明，适合于描述静止、轴对称的黑洞的度规可写成以下的形式：

$$ds^2=\sqrt{(\Delta\delta)}\left[\chi^{(\mathrm{d}t)^2}-\frac{1}{\chi}\left(d\varphi-\omega\mathrm{d}t\right)^2\right]-e^{u_3+u_2}\sqrt{\Delta}\left[\frac{(d\eta)^2}{\Delta}+\frac{(d\mu)^2}{\delta}\right]\quad(1)$$

这里

$$\Delta=\eta^2-1,\ \delta=1-\mu^2=\sin^2\theta\left(\mu=\cos\theta\right)\quad②$$

η是径向坐标（以适当的单位测量），χ、ω、和μ_2 +μ_3 是要确定的度规函数。需要指出的是，ω与黑洞的角动量直接相关；对于静态的史瓦西黑洞，ω等于零。

在写出方程（1）的度规形式中，我们已经考虑到η=1处出现一个零表面，它最终和黑洞的视界相同。

广义相对论的中心问题是解出χ和ω。一旦解出χ和ω，剩下的度规函数$\mu_2 + \mu_3$ 则可用一个简单的面积分求出。

通过变换，我们可以得到度规（1）的一个“共轭度规”，

$$t \to +\mathrm{i}\phi \text{和} \phi \to -\mathrm{i}t \qquad ③$$

通过这种“共轭变换”，χ和ω可以用下面式子替代

$$\tilde{\chi} = \frac{\chi}{\chi^2 - \omega^2} \text{和} \tilde{\omega} = \frac{\omega}{\chi^2 - \omega^2} \qquad ④$$

为了解决这一物理问题，最关键的是一对函数Ψ和Φ取代χ和ω，这里

$$\psi = \frac{\sqrt{\Delta\delta}}{\chi} \qquad ⑤$$

φ是ω的势函数，定义为

$$\Phi_{,\eta} = \frac{\delta}{\chi^2}\omega_{,\mu} \text{及} \Phi_{,\mu} = -\frac{\Delta}{\chi^2}\omega_{,\eta} \qquad ⑥$$

类似地可以根据$\tilde{\chi}$ 和$\tilde{\omega}$ 来定义$\tilde{\Psi}$ 和$\tilde{\Phi}$ 。

在黑洞的数学理论中，把函数Ψ和Φ以及$\tilde{\Psi}$ 和$\tilde{\Phi}$ 组合成一对复函

数：

$$Z^{+}=\Psi+\mathrm{i}\Phi,\tilde{Z}^{+}=\tilde{\Psi}+\mathrm{i}\tilde{\Phi} \quad ⑺$$

并定义

$$E^{+}=\frac{Z^{+}-1}{Z^{+}+1};\ \tilde{E}^{+}=\frac{\tilde{Z}^{+}-1}{\tilde{Z}^{+}+1} \quad ⑻$$

这两个函数满足恩斯特（Ernst）方程

$$\left(1-|E|^{2}\right)\left\{\left[\left(1-\eta^{2}\right)E_{,\eta}\right]_{,\eta}-\left[\left(1-\mu^{2}\right)E_{,\mu}\right]_{,\mu}\right\}=$$

$$-2E^{*}\left[\left(1-\eta^{2}\right)\left(E_{,\eta}\right)^{2}-\left(1-\mu^{2}\right)\left(E_{,\eta}\right)^{2}\right] \quad ⑼$$

接下来我们看看适用于描述撞击波的时空。我们设想两列波前是平面的脉冲引力波，由+∞和-∞处相接近，并发生碰撞（一般说来，上述引力波伴随具有同样波前的引力冲击波和其他冲击波）。在碰撞前瞬间，相互逼近的波前之间的时空是平直的。我们原则上关心的是碰撞后的瞬间时空是如何变化的（尽管不可忽略它们满足相碰波前的边界条件）。

碰撞后的瞬间，时空度规可以写成：

$$ds^2=-\sqrt{(\Delta\delta)}\left[\chi\left(dx^2\right)^2+\frac{1}{\chi}\left(dx^1-q_2dx^2\right)^2\right]+e^{\nu+u_3}\sqrt{\Delta}\left[\frac{(d\eta)^2}{\Delta}-\frac{(d\mu)^2}{\delta}\right] \tag{10}$$

现在这里的

$$\Delta=1-\eta^2,\delta=1-\mu^2 \tag{11}$$

η是碰撞瞬间的时间测度（用适当的单位），μ是碰撞时刻与波前垂直的距离测度，χ、q_2 和$\nu+\mu_3$ 是待定的度规函数。需要注意的是，q_2 直接与引力波极化平面的变化有关。当极化平面保持不变时q_2 为零。

在写出方程（10）的度规形式中，凭经验我们已经考虑到这样一个事实，作为一个碰撞的结果，当η=1且μ=±1时，曲率或坐标奇点开始显示出来。

与静止、轴对称的时空情形相似，只要解出度规函数χ和q_2 ，或解出Ψ和Φ，我们就能够完全解出爱因斯坦场方程，其中Ψ、Φ与χ、q_2 的关系为：

$$\Psi=\frac{\sqrt{(\Delta\delta)}}{\chi} \tag{12}$$

$$\Phi_{,\eta}=\frac{\delta}{\chi^2}q_{2,\mu}\quad \Phi_{,\mu}=\frac{\Delta}{\chi^2}q_{2,\eta} \tag{13}$$

在这种情形下，我们不需要考虑“共轭”过程，因为它对应于χ^1 和χ^2 的一个简单对换。

现在，我们将χ和q_2 以及Ψ和Φ组合成一对复合函数

$$Z=\chi+\mathrm{i}q_2 \qquad Z^+=\Psi+\mathrm{i}\Phi \tag{14}$$

并定义

$$E=\frac{Z-1}{Z+1}, E^+=\frac{Z^+-1}{Z^++1} \tag{15}$$

我们发现E和E+两者都满足同样的恩斯特方程（9）。

当我们考虑带电的黑洞或伴随有电磁波的引力波的碰撞时，我们必须把麦克斯韦方程组补充到爱因斯坦方程组中来。对于即将考虑的对称时空，麦克斯韦场可以用单一的复合势H来表示，有关这个问题的整个方程组最终可以约化为这样一对耦合方程：

$$H\text{和}Z^+=\Psi+\mathrm{i}\Phi+|H|^2 \tag{16}$$

这里Ψ由方程（5）和（12）定义，Φ是ω或q_2 的势函数，后者类似于方程（6）和（13）的定义，但在右边包括了附加项H。

有两种情况可以把影响Z^+ 和H的一对方程约化为一个恩斯特方程，即

$$\text{情形(i)}: H=Q(Z^+ +1) \tag{17}$$

这是Q是某一实常数；还有

$$\text{情形（ii）：} Z=1, \Phi=0\text{和}\Psi=1-|H|^2 \qquad (18)$$

在情形（i）中根据定义

$$E^{+}=\frac{Z^{+}-1}{Z^{+}+1} \qquad (19)$$

我们发现对于我们现在考虑的两类时空，E^{+} 满足方程

$$\left(1-4Q^{2}-|E|^{2}\right)\left\{\left[\left(1-\eta^{2}\right)E_{,\eta}\right]_{,\eta}-\left[\left(1-\mu^{2}\right)E_{,\mu}\right]_{,\mu}\right\}=$$

$$-2E^{*}\left[\left(1-\eta^{2}\right)\left(E_{,\eta}\right)^{2}-\left(1-\mu^{2}\right)\left(E_{,\mu}\right)^{2}\right] \qquad (20)$$

此外，可以证明，如果E_{vac} 是恩斯特方程（9）在真空下的解，那么

$$E_{\mathrm{Ei,Ma}}=E_{\mathrm{vac}}\sqrt{1-4Q^{2}} \qquad (21)$$

则是适合于爱因斯坦—麦克斯韦方程组的方程（20）的解。（需要注意的是：在静止、轴对称情况下，当带有“～”号的变量满足同样的恩斯特方程时，我们也要考虑“共轭”过程）。

在情形（ii）中，我们发现，在真空情况下，当H满足恩斯特方程（9）时，我们有

$$H = E_{\mathrm{vac}} \qquad \Psi = 1 - \left|E_{\mathrm{vac}}\right|^2 \tag{22}$$

要想完全解决我们将考虑的各种问题，特别是撞击波的理论，常常需要相当详细精密的分析。我们不打算讲这样的分析；因为并不需要把整个理论的结构和相关性都陈列出来。

Ⅷ

黑洞和撞击波的数学理论在结构上的相似性，起因于在两种情况下的爱因斯坦和爱因斯坦—麦克斯韦方程能够简化为同样的恩斯特方程；我们将看到，它们甚至有相同的解。这种同一性，只要作特殊的坐标选择，即要保证黑洞的径向距离η=1时出现视界，以及对相互撞击波的时间η=1时能形成—个奇点，我们就能得到这种同一性。尽管存在这种同一性，但对这种物理情况的描绘仍是丰富多彩的，这是与恩斯特方程相同解有关的度规函数的不同组合而产生的结果。

我们将首先考虑从真空方程得到的解。恩斯特方程（9）的解可以描述形形色色的物理现象，它的解是最简单的，即，

$$E = p\eta + \mathrm{i}q\mu \tag{23}$$

这里p和q是两个实常数，它们满足条件

$$p^2 + q^2 = 1 \tag{24}$$

在黑洞理论中，pη+iqμ是 $\tilde{E}^+$ 的解。也就是共轭度规 $\tilde{\Psi} + \mathrm{i}\tilde{\Phi}$ 的恩斯特方程的解，即克尔解。当p=1，q=0时，它就约化为史瓦西解。史瓦西时

空和克尔黑洞在教科书中已经有了充分的描述，并已广为知晓。我将仅仅讨论属予特殊代数类型的解，即彼德罗夫（Petrov）分类中的类型D。属于这种类型的解具有许多特殊的性质，在克尔几何中所有数学物理标准方程的可分离性可归功于这些特殊的性质。

现在谈撞击波理论，其基本解是克罕和彭罗斯（1971）解，它描述了平行极化的两列纯脉冲引力波的碰撞。它得自于恩斯特方程关于$\chi+iq_2$的解E=η。由E=qη+ipμ，人们得到努特库—哈利尔（Nutku-Halil）解（1977），它是描述非平行极化脉冲波碰撞这一更一般情况下的解。这样，克罕—彭罗斯和努特库—哈利尔解在撞击波理论中扮演了和黑洞理论中史瓦西和克尔解相同的角色。

一些度规函数Ψ+iΦ的联立“也会导致同样的恩斯特方程，这引起我们讨论E^+ 的解的pη+iqμ兴趣。下面的解具有我们完全没有预料到的特征：当η=1时，没有出现曲率奇点，却出现了视界，这和通常两列波相碰时类空曲率奇点规则相违背。因此，在这个例子中，我们必须把时空扩展到视界（η=1）以外。做了这样的扩展以后，我们发现扩展后的时空包含了一个域，这个域是落在后面的一个域的镜像；还包含一个未来的域，它含有双曲弧状的奇点。这使我们联想起克尔黑洞内部的环状奇点。值得注意的是，引力波碰撞引起的时空与爱丽思（Alice）在《镜里世界》[19] 的预言极为相似：“房子里装满镜子的走廊很像一眼望不到边的走廊，只是在走廊外就有点不同。”

以上论述中，从E^+ =pη+iqμ得到的解；只适用于当q≠0的情况。当q=0和p=1时，在η=1处产生一个类空曲率奇点，这时时空不能被扩展到将来。

最后，需要指出的是，从E^+ =pη+iqμ，得到的解属于类型D，它具有所有属于这一类型时空的数学特征。

接下来看看爱因斯坦—麦克斯韦方程组，当我们考虑适合于E^+（E^+ 与方程（21）相一致）的恩斯特方程（20）的解E=pη+iqμ（$p^2+q^2=1-4Q^2$）时，我们得到适用于带电黑洞的解。当q=0时，我们得到赖斯纳—诺德斯特姆（Reissner-Nordstrom）解，当q≠0时，我们得到克尔—纽曼解。

在得到撞击波的爱因斯坦—麦克斯韦方程的相应的“基本解”中，存在着概念性的困难。彭罗斯提出过这样一个问题：在外尔张量中，与δ函数奇点相关的引力脉冲波是否意味着在麦克斯韦张量中有一个相似的δ函数奇点？如果是这样，那么电磁场变量的表述中就会含有δ函数的平方根，“人们对怎样解释这样的函数将不知所措”。另外，在满足许多变化的零边界的边界条件问题上，存在着难以克服的困难。由于这些原因，想获得与精心描述初始条件相适应的解的所有努力均告失败。然而，当人们认识到克罕—彭罗斯和努特库—哈利尔解是由恩斯特方程对 $\chi+iq_2$ 的最简单的解得到时，人们很自然会去寻找爱因斯坦—麦克斯韦方程的解。当麦克斯韦场去掉以后，这个解将约化为努特库哈利尔解。但问题并不那么顺当，因为在爱因斯坦—麦克斯韦方程的框架中，在度规函数χ和q_2 的层次上，我们没有一个恩斯特方程。现在已成功地克服技术方面的问题，并能得到一个满足所有必需边界条件和物理要求的解。通过这种“倒转过来的过程”，我们能够获得一个物理上协调的解，这表明广义相对论具有坚实的美学基础。

既然我们有了一个适于$\Psi+i\Phi+|H|^2$ 的恩斯特方程，我们可以认为$E=p\eta+iq\mu$（$p^2+q^2=1-4Q^2$）是适于E+的恩斯特方程（20）的解。当Q=0时，这个解就约化为我们先前讨论过的真空解，我们发现，像真空解一样，这个解会出现一个视界，和继之产生的类时奇点。

在第Ⅶ节讨论爱因斯坦—麦克斯韦方程时，我们区分了两种情形即情形（i）和情形（ii），它们是有根本的区别。当电磁场去掉以后，情形（i）中的时空就约化为爱因斯坦真空方程的特殊解；而在情形（ii）中，时空变得平坦起来。至今为止，我们所讨论的解都属于情形（i）。我们知道，在情形（ii）中，复合电磁势H满足真空的恩斯特方程（9）。因此我们自然要关心恩斯特方程最简单的解$p\eta+iq\mu$的时空特征。对此所求得的解是很值得人们注意的（由贝尔和切克尔斯（Bell and Szekeres）用不同的方法得到）：用非零的外尔标量表明的引力，完全被约束在描述脉冲引力波的δ函数范围以内。换句话说，如果没有脉冲波的出现，时空是相当平坦的。这样，作为爱因斯坦—麦克斯韦方程的精确解，我们就可以得到一个相似的平坦时空，在这种时空里，伴随脉冲引力波与具有平面波前的电磁冲击波的相互撞击，并产生一个视界。贝尔—切克尔斯解的另一个特征是q=0时，这个解完全等价于q≠0时的解。因此，为了在描述物理状况方面得到一个比贝尔—切克尔斯解更

一般的解，我们应该考察恩斯特方程最简单的解的范围之外的情况。为此，我们利用埃勒斯（Ehlers）变换，可以从任何给定的恩斯特方程解中得到一个单参量的解系。因此，我们考虑解E=pη+iqμ的埃勒斯变换。我们发现，最后所得的解属于类型D，并具有从解E=pη+iqμ得出的真空解的全部特征。特别值得注意的是，通过将埃勒斯变换应用到贝尔—切克尔斯解，我们将得到一个完全具有以上结构的单参量的时空簇。

在表1中，我们很全面地列出了由黑洞和撞击波方程得到的各种各样的解。从表中可以清楚地看出在结构上它们具有统一性。

黑洞理论和撞击波理论之间的内在关系，可以从较简单的情形（见表2），即当ω=0和q_2 =0时清楚地看出。在这种情形下，两种理论赖以求解的基本方程为：

$$\left[\left(1-\eta^{2}\right)\left(\ln\Psi\right)_{,\eta}\right]_{,\eta}-\left[\left(1-\mu^{2}\right)\left(\ln\Psi\right)_{,\mu}\right]_{,\mu}=0 \qquad (25)$$

这个方程可以精确求解，与两个理论相关的解列于表2。

以上讨论表明：爱因斯—坦麦克斯韦方程具有许多与爱因斯坦真空方程相同的显著特征。除麦克斯韦场外，只有一种源能够与引力耦合导致至少保留真空方程的某些特征，这种源是一种满足状态方程的理想流体，能量密度（ε）=压强（p）。对这一理想流体，根据爱因斯坦方程，里奇（Ricci）张量由下式给出

$$R^{ij}=-4\varepsilon u^{i}u^{j} \qquad （26）$$

这里u^i 代表流体的四元速度。

假设在撞击波相互作用范围内，撞击发生后的瞬间，我们可以得到一个具有ε=p的理想流体作为源，我们发现，在撞击瞬间之前，脉冲的引力波必然伴随具有以下动量张量的零尘（nulldust）：

$$T^{ij} = Ek^i k^j = -\frac{1}{2} R^{ij} \qquad (27)$$

这里E是某种正标量函数，k^i 代表一个零矢量。换句话说，在这种假想的情况下，在碰撞的一瞬间，零尘（即描述零轨迹无质量粒子）转变成一理想流体（它的流线沿着类时轨迹）。在碰撞的一瞬间，这种转变是令人惊奇的。但是，正像罗杰·彭罗斯和李·林德布卢姆（Lee Lindblom）指出的那样，上述转换在狭义相对论框架中也能发生，只是这一事实以前似乎并没有注意到。

通过与黑洞理论相比较来阐述撞击波理论时，我们事实上系统地考察了在各种情况下采用恩斯特方程最简单解的结果（在埃勒斯变换情况下）。虽然这种探讨方式完全是形式上的，但还是揭示了人们以任何方法都不能预见的一些可能性，比如：视界、类时奇点或从零尘到理想流体的转变等的出现。在这种情况下，从广义相对论美学基础的角度探讨，可以加深我们理解这一理论的物理意义的深度。

1915年11月爱因斯坦首次公布了他的场方程，在文章结尾处他说：

> 任何充分理解这个理论的人，都无法逃避它的魔力。

至少对于一个多年从事研究的科学家来说，这个理论的魔力就在于它的数学结构和谐一致、前后贯通。

附录

1916年6月29日，在柏林科学院的会议上，爱因斯坦做了一个纪念卡尔·史瓦西的简短致词，对他进行了比较恰当的评价。我想有必要将爱因斯坦的致词翻译在下面。

1916年5月11日，死神从我们队伍中夺去了卡尔·史瓦西，他仅仅活了42岁。这位具有极高天赋、学问渊博的科学家的夭折，不仅是我们科学院，而且也是天文学界和物理学界所有朋友们的悲痛和损失。

在史瓦西的理论工作中，特别使人感到惊讶的是他那么有把握地运用数学研究方法，轻捷地理解天文学或物理学问题实质的本领。很少人具有像他这样把正确的想法和思维的灵活性相结合的深刻的数学知识。正是有了这些才能，才使他在别的研究工作者被数学困难吓住了的那些领域中，完成重要的理论工作。显然，他的源源不断的创作动力，在更大的程度上可以认为是数学概念之间精美联系的那种艺术家的喜悦，而不是要去认识自然界中尚未发现的关系的渴望。因此，可以理解为什么他最初的理论工作属于天体力学，这个知识部门的基础比起其他任何精密科学部门的基础来，可以在更大程度上认为是已经完全建立起来的。在这些论文中，我在这里要提到的，只是关于三体问题周期解的论文，以及关于庞加莱的转动流体平衡理论的论文。

史瓦西最重要的天文学论文是他关于星体统计学的研究。星体统计学是一门试图按照那些包括我们的太阳在内的恒星亮度、速度和光谱类型的观测资料的统计规律，来确定这些巨大的天体构造的科学。在这个领域内，天文学界靠他来进一步深化和发展了对卡普坦新发现的了解。

史瓦西用他在理论物理学方面深刻的知识来构建关于太阳的理论。在这方面，他关于太阳大气的力学平衡和太阳光辐射的测定过程的研究，博得了科学家们的赞扬。在这里应该提到他关于光给小球体压力的优美论文，它使哈雷彗星尾部理论能够建立在牢固的基础上。尽管这项理论研究是用来解决天文学问题的，但它也表明，史瓦西感兴趣的范围也包括纯物理学问题。由于他对电动力学基础有价值的研究，我们应当

给予他应有的赞扬。在他一生的最后一年中，他还对新的引力理论十分感兴趣，而且他首先在新引力理论中得到了精确的计算。在他一生的最后几个月里，病魔已经开始消耗他的体力，他还是成功地在量子论某些方面完成了意义深远的研究。

在史瓦西的伟大理论贡献中，还包括了他关于几何光学的研究，在这些研究中，他改进了对天文学十分重要的光学仪器的误差理论。仅仅这些使仪器得到改进的成果，就已经足以看出他对这门科学的巨大贡献了。

史瓦西的理论工作与他经常的天文学实践活动是紧密结合的。他自24岁以来，就一直未中断在天文台工作：1896—1899年任维恩的助手；1901—1909年任哥廷根大学天文台台长；1909年以后任波茨坦天体物理研究所所长。作为天文观测者和领导人的活动，在他一系列的研究中得到了反映。由于他精力充沛，他还在他的研究领域内制定了新的观测方法。在实验物理中，他借助于照相底片变黑的规律，用照相方法来达到测量光度的目的。为了纪念他，人们把这种测量方法以他的名字命名。他还天才地利用星体的焦外象来测量星体的亮度。由此，除了肉眼观测以外，只有借助他的这一思想，摄影光度学才得以建立。

自1912年以后，由于他在短短时间内取得的成就，这位谦虚的人成了柏林科学院院士。无可挽回的死亡带走了他，然而他的工作是卓有成效的，并给他贡献了毕生精力的科学带来了深远的影响。

表**1**

关键向量	场方程	恩斯特方程解			解	描述对象
		E	E^+	$\tilde{E}^+$		
$\partial_t, \partial_\phi$	爱因斯坦 真空方程	不存在		η	史瓦西	黑洞；静态；球对称 视界 中心的类空奇点 类型D；参量：质量
$\partial_t, \partial_\phi$	爱因斯坦 真空方程	不存在		$p\eta+\mathrm{i}q\mu$ $p^2+q^2=1$	克　尔	黑洞；静态；轴对称 视界和柯西视界；能层 赤道平面上的类时环状奇点 类型D；参量：质量和角动量
$\partial_t, \partial_\phi$	爱因斯坦— 麦克斯韦方程	不存在		$\eta\sqrt{(1-4Q^2)}$	赖斯纳—努德斯特姆	带电黑洞；静态；球对称 视界和柯西视界 中心的类时奇点 类型D；参量：质量和电荷
$\partial_t, \partial_\phi$	爱因斯坦— 麦克斯韦方程	不存在		$p\eta+\mathrm{i}q\mu$ $p^2+q^2=$ $1-4Q^2$	克尔—纽曼	带电黑洞；静态；轴对称 视界和柯西视界；能层 赤道平面上的类时环状奇点 类型D；参量：质量、电荷和角动量
∂_χ^1，∂_χ^2	爱因斯坦 真空方程	η		χ^1和χ^2互换	克罕—彭罗斯	脉冲引力波的碰撞 平行极化 出现类空曲率奇点
∂_χ^1，∂_χ^2	爱因斯坦 真空方程	$p\eta+\mathrm{i}q\mu$ $p^2+q^2=1$		χ^1和χ^2互换	努特库—哈利尔	脉冲引力波的碰撞 非平行极化 出现类空曲率奇点（比克罕彭—罗斯弱）

续表1

关键向量	场方程	恩斯特方程解			解	描述对象
		E	E^{+}	$\tilde{E}^{+}$		
∂_{χ}^{1}，∂_{χ}^{2}	爱因斯坦—麦克斯韦方程	不存在	$E^{+}(E_{vac}=p\eta+q\mu$ $\mathrm{i}q\mu)\times\sqrt{(1-4Q^{2})}$	χ^{1}和χ^{2}互换	钱德拉塞卡—克桑索普洛斯	伴随引力波和电磁冲击波的 脉冲引力波的碰撞 非平行极化 出现类空曲率奇点
∂_{χ}^{1}，∂_{χ}^{2}	爱因斯坦真空方程		η	χ^{1}和χ^{2}互换	钱德拉塞卡—克桑索普洛斯	伴随引力冲击波的脉冲引力波的碰撞 平行极化 出现非常强的类空曲率奇点类型D
∂_{χ}^{1}，∂_{χ}^{2}	爱因斯坦真空方程		$p\eta+\mathrm{i}q\mu$ $p^{2}+q^{2}=1$	χ^{1}和χ^{2}互换	钱德拉塞卡—克桑索普洛斯	随引力冲击波的脉冲引力波的碰撞 非平行极化 出现一个视界和随后的类时奇点类型D
∂_{χ}^{1}，∂_{χ}^{2}	爱因斯坦—麦克斯韦方程	不存在	$\eta\sqrt{(1-4Q^{2})}$	χ^{1}和χ^{2}互换	钱德拉塞卡—克桑索普洛斯	伴随引力波和电磁冲击波的脉冲引力波的碰撞 平行极化 出现一个视界和随后的三维类时奇点类型D
∂_{χ}^{1}，∂_{χ}^{2}	爱因斯坦—麦克斯韦方程	不存在	$p\eta+\mathrm{i}q\mu$ $p^{2}+q^{2}=$ $1-4Q^{2}$	χ^{1}和χ^{2}互换	钱德拉塞卡—克桑索普洛斯	伴随引力波和电磁冲击波的脉冲引力波的碰撞 非平行极化 出现视界和随后的类时弧状奇点类型D

续表**2**

关键向量	场方程	恩斯特方程解			解	描述对象
		E	E^+	$\tilde{E}^+$		
∂_χ^1，∂_χ^2	爱因斯坦—麦克斯韦方程（$H=E_{vac}$）	不存在	η	χ^1和χ^2互换	贝尔—切克尔斯	伴随电磁冲击波的脉冲引力波的碰撞 平行极化 时空平坦 出现视界；允许没有随后奇点的延拓
∂_χ^1，∂_χ^2	爱因斯坦—麦克斯韦方程（$H=E_{vac}$）	不存在	$p\eta+iq\mu$ $p^2+q^2=1$	χ^1和χ^2互换	贝尔—切克尔斯	同上
∂_χ^1，∂_χ^2	爱因斯坦—麦克斯韦方程（$H=E_{vac}$）	不存在	$p\eta+iq\mu$的欧勒变换	χ^1和χ^2互换	钱德拉塞卡—克桑索普洛斯	随引力波和电磁冲击波的脉冲引力波的碰撞 出现视界和随后类时弧状奇点 类型D
∂_χ^1，∂_χ^2	爱因斯坦—流体动力学方程（$\varepsilon=p$）	$p\eta+iq\mu$		χ^1和χ^2互换	钱德拉塞卡—克桑索普洛斯	伴随有引力冲击波和零尘（$R_{ij}=Ck_ik_j$）的脉冲引力波的碰撞 非平行极化 出现弱类空奇点 零尘转变成理想流体$\varepsilon=p$

表**2**

关键向量	场方程	解	评注
∂_t,∂_ϕ	爱因斯坦 真空方程	$\ln\Psi=\ln\frac{\eta-1}{\eta+1}$	史瓦西解：静态球对称黑洞
∂_t,∂_ϕ	爱因斯坦 真空方程	$\ln\Psi=\ln\frac{\eta-1}{\eta+1}+\sum_n A_nP_n(\mu)P_n(\eta)$	畸变黑洞（当$\sum A_{2n+1}P_{2n+1}(1)=0$时）；外尔解
∂_χ^1，∂_χ^2	爱因斯坦 真空方程	$\ln\chi=\ln\frac{1+\eta}{1-\eta}$	平行极化的脉冲波碰撞的克罕—彭罗斯解
∂_χ^1，∂_χ^2	爱因斯坦 真空方程	$\ln\chi=\ln\frac{1+\eta}{1-\eta}+\sum_n A_np_n(\mu)p_n(\eta)$	伴随引力冲击波的脉冲引力波的碰撞； 平行极化
∂_χ^1，∂_χ^2	爱因斯坦— 麦克斯韦方程 (H=E_{vac})	$\frac{1}{2}\ln\frac{1+H}{1-H}=\frac{1}{2}\ln\frac{1+\eta}{1-\eta}$	平坦时空中伴随电磁冲击波的脉冲引力波的碰撞；平行极化
∂_χ^1，∂_χ^2	爱因斯坦— 麦克斯韦方程	$\frac{1}{2}\ln\frac{1+H}{1-H}=\frac{1}{2}\ln\frac{1+\eta}{1-\eta}+\sum_n A_np_n(\mu)p_n(\eta)$	伴随有引力波和电磁冲击波的脉冲引力波的碰撞；平行极化

注：基本方程，$\left[(1-\eta^2)(\ln\Psi)_{,\eta}\right]_{,\eta}-\left[(1-\mu^2)(\ln\Psi)_{,\mu}\right]_{,\mu}=0$

附录　寻求秩序——钱德拉塞卡对黑洞、蓝天和科学创造力的思考

John Tierney

钱德拉塞卡也不能肯定他自己下一步将做什么研究。他最近是在从事黑洞的研究，这项研究已进行了8年。在这期间，他的心脏病犯了，为此他做了体外循环的心脏手术。这年春天他完成了黑洞的研究，这时他已有71岁的高龄，几乎是这个领域里其他同行年龄的两倍。大多数科学家在这个年龄，要么退休，要么享受一些名誉头衔——在一些委员会里担任职务、在颁奖宴会上回忆往事、指导几个研究生、琢磨几个尚未解决的问题使自己得到愉悦，等等。但这对钱德拉塞卡来说，几乎是不可想象的。要他这样打发日子，无异于要他上班时不穿白衬衫、黑西服，不打黑领带。在近半个世纪里，他总是这样正儿八经地穿戴整齐地到芝加哥大学上班。

的确如此！每当他投入工作时，他就会坐在一张非常整齐、清洁的书桌前，寻觅数学的秩序；而且每天至少工作12小时，一周工作7天。如此勤奋工作约10年时间，当他得到了“某种见解”以后才罢休——也就是说，直到宇宙的某一个方面已经完全约化为一组方程时才罢休。然后，把他研究的结果写在一本书里，就不再注意这个研究课题，而去寻找天体物理学领域里另一个完全不同的课题，重新埋头研究下去。这种研究风格被人们称为“钱德拉风格”。这种风格简直让其他天文学家感到头晕眼花，无法理解。在天体物理学领域里，一般人认为40岁的理论家都已经远远超过了做出重要发现的黄金年龄，但年已63岁的钱德拉塞卡却会强迫自己放弃已经驾轻就熟的课题，重新开始分析当物质消失在一个黑洞里时，会发生什么现象！这真令人难以理解。

他最亲近的朋友，普林斯顿大学的天体物理学家史瓦西（M.Schwarzschild）说：“他那种不怕困难的奋斗精神，是别人做不到的。钱德拉的专心致志的能力难以令人相信，这是一种真正的数学才智

和惊人毅力的结合。在他所研究的任何一个领域里，他都得出了一些我们现在天天在应用的重要结果。”

钱德拉塞卡常常喜欢从哲理方面谈创造力和科学思想老化的问题，而且谈起来总是入木三分。但是，每当问及他的研究生涯何以能持续如此长久时，他往往不自在起来。但他承认，1935年1月11日在（英国）皇家天文学会会议上发生的事情，也许可以发现一些端倪。

那天是星期五，他怀着极大的期望，同时又夹杂着对爱丁顿爵士的某些疑虑，来到了伦敦。最近几个月里，他和爱丁顿大约每周两次晚饭后在一起，共同讨论钱德拉塞卡最近对正在衰亡的恒星行为所作的计算。他们俩是一对奇怪的组合：爱丁顿52岁，风度翩翩，说话很有吸引力，他被认为是世界上最优秀的天文学家；而钱德拉塞卡是一位腼腆的来自印度的24岁的青年，自认为有点像剑桥大学的流浪儿。钱德拉塞卡从事星体结构的研究只有几年时间，那还是起因于他在马德拉斯（Madras）大学的一次物理竞赛中获奖，奖品是爱丁顿写的一本关于星体结构的论著。现在，他相信自己有了一个重要而且惊人的发现，这个发现将在星期五的会议上向与会者宣布。

但是在星期四，当会议日程表送到剑桥时，钱德拉塞卡惊讶地发现，爱丁顿也将在会议上发言，讲的题目与他的一样！在他们多次讨论过程中，钱德拉塞卡总是滔滔不绝地讲出他的计算，而爱丁顿从来没有提到过他在这个领域的工作。这似乎是一种令人难以置信的不忠诚行为。星期四晚上，当他们两人在餐厅相遇时，爱丁顿仍然没有提出解释或道歉。他唯一的话是，他利用了他的影响，使钱德拉塞卡可以在会议上多讲一会，“这样，你就可以把你的研究结果彻底讲清楚。”这句关心的话钱德拉塞卡记得很清楚。钱德拉塞卡想问一问爱丁顿的论文，但由于他对爱丁顿非常尊敬，所以又不敢问，星期五在伦敦开会前吃茶点时，有一位天文学家问爱丁顿准备讲些什么，爱丁顿没有回答，只是转向钱德拉塞卡，微笑地说：“那要使你大吃一惊呢。”

钱德拉塞卡的论文讨论了一个基本问题：一个星体如果烧尽了它全部的燃料之后，将会发生什么？按照当时流行的理论，冷却的星体将在自身的引力作用下，坍缩为一颗致密的星，这种星称为白矮星（white dwarf）。例如，一颗具有太阳质量的星，将坍缩成只有地球一样的大小，这时它就达到力的平衡，不再坍缩下去。钱德拉塞卡重新研究了这

种坍缩，他考虑一个有趣的问题：当一颗星体的气体被压缩到其电子以接近光速运动时（这被称之为相对论性简并状态），会发生什么现象？他的研究结果表明，任何一颗质量大于太阳质量1.4倍的星体，它巨大的引力将继续起作用，使这颗星越过白矮星阶段继续坍缩。于是，这颗星的体积将继续越变越小、密度越来越大，直到……啊，那可是一个有趣的问题。钱德拉塞卡机灵地避不作答。

他的结论是："一颗质量大的星体不会停留在白矮星阶段，我们应该考虑其他的可能性。"

接着是爱丁顿发言。

"我不知道我是否能活着逃离这个会议，但我的论文的要点是：根本不存在什么相对论性简并。"爱丁顿说完这句话，就把钱德拉塞卡的论文撕成两半。听众中不时爆发出笑声，使爱丁顿的讲话不断地被打断。爱丁顿无法驳倒钱德拉塞卡的逻辑和计算，但他却宣称这个理论全盘皆错，原因很简单，因为它得出了一个不可避免和非常古怪的结论："星体将不断辐射，不断收缩。直到半径只剩下几公里，这时，引力将大到足够控制住辐射，使其不再辐射，于是这颗星体终于平静下来。"

当然啦，我们今天称这种客体为黑洞，但爱丁顿那天下午却声称它不可能存在。

"一种归谬法得出的结论，"爱丁顿说，"我相信，一定有一条自然法则阻止星体按这种荒谬的方法演化。"

争论到此，就被束之高阁了，至少在此后几十年没人再敢问津。当然，黑洞理论最终还是被人们承认了，钱德拉塞卡提出质量界线（即星体质量为太阳质量的1.4倍）也被称之为"钱德拉塞卡极限"记入教科书中，但这一切在爱丁规讲话后很长一段时间才发生。

"会议结束后，"钱德拉塞卡回忆说，"每个人都走到我面前说，'太糟了，钱德拉，简直太糟了。'我原以为我将在会上宣布一个非常重大的发现，结果爱丁顿让我出足了洋相。我心情乱极了。我不知道我是否应该继续研究这个课题。那天我回到剑桥时，已是深夜一点多了。我走进一间同事们经常聚会的房间，这时当然不会有人还留在那里，但炉子

里的火还燃烧着。我记得，我站在炉火前自言自语地重复着一句话：‘世界就是这样结束的，不是伴着一声巨响，而是伴着一声呜咽。’”

今天，他对那天下午发生的事有了不同的看法。

与爱丁顿的争论还持续了几年时间，这使钱德拉塞卡失去了在英国取得一个有任期职位的任何机会。最后，他终于明白应该完全放弃这个研究课题（不过，这两位都十分出色的人终生都保持着友谊）。他相信自己的理论，但别人不相信。所以，在1937年他到了芝加哥大学以后不久，他把他的理论写进了一本书里，并不再去理会它。他改弦易辙，开始研究星体在星系中的几率分布，并发现了被称之为动力学摩擦的奇怪特性——即任何星体在穿越另外一个星系时，由于它四周星体的引力作用，它的速度将会降低。然后，他又转而考虑：天为什么是蓝色的？这个问题的简单答案是：大气的分子允许其他颜色的光通过，但将波长短的蓝色光散射开。这个结论在上一个世纪被英国的瑞利爵士发现，但瑞利本人及后来的物理学家们没有能够找到精确的数学方法，以计算光是如何被散射的。在本世纪40年代中期，钱德拉塞卡找到了这个精确的数学方法。对此，钱德拉褰卡十分满意，以至他决定此后将不断地更换研究领域。接下去他研究了许多课题：磁场中热流体的行为，旋转物体的稳定性，广义相对论，最后，他又回到了黑洞，但这次他从一种完全不同的角度研究它。他现在感到十分幸运，幸亏当年他被赶出原来研究的专业。

“假定当时爱丁顿同意自然界有黑洞，”说到这儿钱德拉塞卡停顿了一下，考虑如何把这个见解表达得尽可能准确。

“推测是非常困难的。如果爱丁顿当时承认了黑洞理论，他将会使这整个领域变成一个引人注目的研究领域，黑洞的许多性质将会提前20到30年被人们发现，不难想象，理论天文学将会大不相同。但这种判断不该由我来作——喏，我想我可以说，这种结局对天文学是有益处的。”

“但我不认为对我个人有益。爱丁顿的赞美之词将使我那时在科学界的地位有根本的改变，我会在天文学界里大有名气。但我的确不知道，在那种诱惑的魔力面前我会怎么样。”

“有多少年轻人在功成名就之后，还能长久保持青春活力呢？即使

是20年代里对量子力学做过伟大贡献的科学家——我指的是狄拉克、海森伯、福勒——他们也未能始终如一。即使是麦克斯韦和爱因斯坦，也同样未能始终如一。”

钱德拉塞卡急忙停住自己的话头，说他并不是要把自己与这些科学家相比，也并不是想批评他们。“你可别搞混了，我算老几，竟能批评爱因斯坦？”他认为，这个问题只是在抽象的意义上使他感到兴趣。贝多芬在47岁时对一位朋友说：“现在我知道怎么作曲了。”这件事使钱德拉塞卡大为吃惊，留下了难忘的印象。钱德拉塞卡相信，决不会有一位年满47岁的科学家会宣称：“现在我知道怎么做研究了。”

“当你讨论一位伟大的艺术家或作家的作品时，总会认定这些作品有一个从早期、中期到成熟期的发展过程。艺术家的本领越来越精，使他们最终能处理十分棘手的难题。要写出一部像《李尔王》这样的剧本，显然需要付出巨大的努力，要具备巨大的感情控制能力。如果与早期作品《罗密欧与朱丽叶》相比较，你就会发现其间的差别。”

“现在我们来看一看，为什么科学家不能使他们的思想越来越精呢？爱因斯坦是一位具有伟大科学头脑的科学家。他在1905年就发现了狭义相对论和一些其他重大理论。他以惊人的刻苦精神从事研究，1916年他又发现了广义相对论；到20年代早期，他还做出过一些重要的发现。但从此以后，他停步不前，孤立于科学进步之外，成为一个量子理论的批评家，再没有为科学和他本人增添什么光彩。爱因斯坦在40岁以后的研究工作中，没有任何迹象表明他的智慧和悟性比以前更强一些。这是为什么？”

“由于缺乏更合宜的词，我只能说，似乎是人们对大自然逐渐产生了一种傲慢的态度。这些人曾经具有伟大的洞察力，做出过意义重大的发现。此后，他们就相信，他们能在一个领域里取得如此辉煌的胜利，说明他们有一种看待科学的特殊方法，而且这种方法一定是正确的。但是，科学并不承认这种不切实际的想法。大自然曾一次又一次地表明，作为大自然基础的各种真理，比最聪明的科学家更加强大和有力。”

“以爱丁顿为例。他是一位伟人，他说一定有一条自然法则阻止星体变成一个黑洞。他怎么会说这样的话呢？仅仅是因为他认为那样不好吗？为什么他能认定，他有办法决定自然法则应该是怎样的呢？相似

地，爱因斯坦因为不赞成量子理论说过一句人人都知道的话：‘上帝不会掷骰子。’他怎么知道上帝不掷骰子呢？”

谈到瑞利爵士，钱德拉塞卡乐于承认他是一个例外。瑞利是19世纪的一位物理学家，他在50年的研究生涯中，在各种不同领域里都做出了富有创造性的贡献，而且在后期还做出了一些最著名的发现，例如发现了氩气。

“你一定听说过，当瑞利67岁时，他的儿子曾经问他对托马斯·赫胥黎说的一句名言有什么看法，赫胥黎说：‘在科学界一个60岁的人的作为只会弊多利少。’瑞利想了很久，回答说，‘啊，我想如果你只做你所理解的事，不和年轻人发生矛盾，就不致于一定会像赫胥黎说的那样。’我认为爱因斯坦、狄拉克或海森伯不会说出这种话来。爱丁顿肯定说不出这种话的。瑞利说的话中，含有某种谦虚的精神。有人曾经向丘吉尔说艾德里是一个谦虚的人，丘吉尔回答说‘他还有许多方面需要谦虚’。丘吉尔的这句话完全可以用到科学界。能够做出真正伟大发现的人，一般都非常自信，敢于对大自然做出判断。当然啦，瑞利不具有爱因斯坦或麦克斯韦那样的真正伟大的根本性的洞察力。但是，瑞利对科学的影响是巨大的，因为他对伟大的知识宝库增加了一些内容，不断地发现了许多虽不是蔚为壮观但却很重要的东西。我认为在理解大自然的过程中具有某种谦虚精神，是持续进行科学探索的先决条件。”

他一再声明，他只是从抽象的意义讨论这个问题，而不是说他自己。但他不可避免地会谈到他自己的研究生涯。每隔十年就投身于一个新的研究领域，可以保证研究者保持谦虚的精神。因为你无法与年轻人发生矛盾，他们在新领域里研究的时间比你还长。像瑞利一样，钱德拉塞卡使自己关心那些重要但并不壮观的工作，以严格的研究精神为知识宝库增添内容，而不是去设法推翻它。他不追求轰动效应，不追求独一无二的令人目眩的洞见，也不追求可以获得诺贝尔奖的革命性发现。他总是坚持对整个领域作长期和全面的分析，而不计较别人认为他的做法可能毫无价值。

例如，在20世纪60年代，钱德拉塞卡写过一本关于蜜柑状的几何图形（称为椭球）的书，这种工作在当时肯定不会给任何人带来名和利。在该书的引言中钱德拉塞卡写道，他写这本书的理由是，以前虽然对这项课题做过许多研究，但留下“许多空白和漏洞，还有一些明显可见的

错误和谬见。如果对这种状态听之任之，不免令人感到遗憾。”他通过系统地分析作用在旋转椭球体上的诸力，例如使它聚在一起的引力、使它分裂的离心力等，终于将这个难题理出了一个头绪，他找到了使旋转椭球体开始变得不稳定的点。其他科学家认为，研究这些理想物体是浪费时间。研究宇宙中不存在的抽象东西有什么价值？但到20年以后的今天，这本书却得到了广泛的应用，这真是人们始料未及的。例如，现在已经知道许多真实的星系具有这些想象客体的性质，科学家们正在利用这本书的结论来了解，银河系在自旋时为什么能聚在一起而不散开。

“我想，我的动机与许多科学家的不一样，”钱德拉塞卡说，“沃森（James Watson）[20]说，当他还是年轻人时，就想到要解决一个问题，使他因此而获得诺贝尔奖。他努力向前干下去，后来发现了DNA。很明显，这种方法对沃森是很合适的，但我的动机不是为了解决一个单一的问题，我需要的是对整个领域有一种透视的看法。”

“8年前我开始研究黑洞，而且特别专注于研究一个旋转的黑洞对外来的（如引力波、电磁波）扰动如何反应。如果把这个问题弄清楚了，你就可以知道当一个诸如星体之类的物体落入黑洞时，会发生什么情况。你瞧，这项研究的某些成果已开始引起人们的注意了，但对我来说，最重要的是研究的最终观点。这就是我为什么写这本书的缘由，我把它看作是能够洞察未来的一个整体。很明显，在这个领域里还有许多工作让我干，但我不想干下去了。打个比方，如果你完成了一个雕塑作品，你一旦完成就不想再在这儿或那儿进一步修饰它。”

下一步做什么呢？对于一个年届71岁的人来说，这可是一个问题。主要的障碍是由他的研究风格决定的，这种风格需要他投入大量的时间和精力，意味着他将每天从早晨六点钟一直干到半夜。

钱德拉塞卡的夫人叫拉莉莎（Lalitha），他们是在马德拉斯大学物理系读书时相识的，她说：“他的确没有时间干别的事，例如旅游、会朋友。他对工作总是制定严格的纪律，干任何事都一定要干得干净利落、完全彻底。”谈到钱德拉塞卡的事业和她为之付出的牺牲，她毫无怨言。她经常孤独一人在家打发日子，她放弃了她的职业跟他来到美国，要隔很多年才有一次机会回印度与亲人团聚。她也认为，钱德拉塞卡现在应该有权力轻松一下，他的同事们也这么认为。

“钱德拉塞卡曾不得不付出巨大的代价，随着他逐渐变老，他所付出的代价也在逐步增加，”史瓦西说，“最近他写的一本书可说是一部力作，也是意志战胜年衰体弱的一个明证。我的确不知道他下一步还会干什么。他所有的贡献都来自于这种坚强的意志力，这使得他能够解决别人无法解决的问题。对钱德拉来说，如果在一件大家都能做的事情上泡蘑菇，那是与他性格不能相容的事情。”

钱德拉塞卡比较同意这种说法。他说：“如果我不能全心投入一项研究项目，我宁可完全放弃这项研究。”如果只是为了美学上的理由，那么他的学术生涯以黑洞理论为开端，又以黑洞理论来结束，实在是颇能吸引人的。这个结局实在太妙了，特别是因为他认为这最后的一项研究是他所有研究工作中最困难的一项。但是，钱德拉塞卡还打算进入其他新的领域，这次也许是宇宙学。“我的生活习惯就是这样的，要改变终生形成的习惯可不那么容易。不过，我还没有拿定主意。”

虽然他不再认为自己是一个印度教徒，而认为自己可以划归为无神论者，但有时他也有点拿不定主意，是不是应该跟随印度教的传统：断绝一切世俗的联系，到森林里一个人去冥思苦想。当然啦，对他来说这就意味着完全放弃科学，最终地变换领域。

“我追求科学的方式给我带来不幸的事实是，它扭曲了我的个性。我不得不放弃生活中其他一些兴趣，例如文学、音乐、旅游等等。我把所有的时间都奉献给了科学事业。我想精读莎士比亚所有的剧本，一字也不漏，一行也不漏。但我总是没有时间去读。如果我能满足我的兴趣，我将成为另外的一个人，这点我十分清楚。我不知道，能不能用遗憾这个词表达我的这种感觉。但人迟早总会习惯这些损失，人总不会跟自己过不去，他总得跟自己妥协。要把事物整理得有条有理，那总是要花费一定的时间。”

（译自Science sept.1982，pp.69-74）

译后记

杨建邺

1994年6月8日于武汉

华中理工大学宁泊书斋

钱德拉塞卡教授是美籍印度天体物理学家，1983年诺贝尔物理学奖获得者，他获奖的原因是“因为对恒星结构和演化过程的研究，特别是因为对白矮星的结构和变化的精确预言”。钱德拉塞卡著作颇丰，除了我们译的这本书以外，还有许多专业性非常强的著作，如《恒星结构研究导论》（1939年）、《恒星动力学原理》（1942年）及《黑洞的数学理论》（1983年）等等。

我是学物理出身的，对钱德拉塞卡传奇般的经历早有所闻，但却从来没有读过他的著作。1989年7月24日，我忽然收到在美国纽约工作的大哥寄来的一包书，打开封皮，我一眼就盯上了钱德拉塞卡著的《真理与美》（Truth and Beauty）。按惯例，我翻开目录：“科学家”、“科学的追求及其动机”、“莎士比亚、牛顿和贝多芬：不同的创造模式”，还有“广义相对论的美学基础”！我似乎觉得眼睛一亮，一个崭新的世界在我面前打开了。一篇一篇看下去，这种感受越来越强烈。正如作者在前言中所说，他思考的是一些我们大家应该思考但又“从未认真思考过的问题”；而且我还深深感到，钱德拉塞卡思考的这些问题对中国读者一定很有价值。于是我决心将这本书译出。开始我在《世界科学》上陆续译出了几篇，颇受编辑和读者的青睐，这更加深了我的信心：一定要把这本书奉献给中国读者。

但几经碰壁之后，我才知道目前在国内想出版一本这类高级科普读物，可以说非常困难。幸运的是湖南科学技术出版社决定将这本译著放到《第一推动丛书》里，这使我大为振奋。

出版社要求我与钱德拉塞卡教授取得联系，解决译文版权问题。这是很公正的要求，作为译者我也应该尊重钱德拉塞卡教授的知识产权和个人意愿。但我非常担心，钱德拉塞卡教授年事已高，今年84岁了，又是蜚声科学界的大人物，他会给我回信吗？但我别无选择。好不容易查到了他的地址，于是在今年3月底给他写出了第一封信，请他允许我将他的这本著作译成中文，并请求他将中译本的版权授予我们。出乎意料的是，不到一个月我就收到了钱德拉塞卡教授的回信，我真的感动极了。在回信中他写道："您想翻译我的书，我当然非常高兴；作为作者，我允许您进行翻译。"

我高兴极了，这不仅是因为我获得钱德拉塞卡教授的允许，可以翻译这本名著，还因为广大中国读者将有幸与这位诺贝尔奖得主一同去思考那些本应认真思考但从未认真思考过的一些重要问题。

闲话不多说了，但有两件事我还得向读者交待一下：一是我把原书名《真理与美》改为现在使用的书名，是为了让更多年轻读者愿意买这本书。我相信，现在使用的书名对广大读者一定更有吸引力；我还相信，我并没有因改了书名而背离了原作者的意图（本书在2016年改版时，编者经考虑将本书名改回了原书名《真理与美》——编者注。）。二是我在书末加了一个附录，这是为了让读者熟悉钱德拉塞卡教授年轻时一段传奇般的经历，了解了这段经历会使读者更深刻地领会作者的思路。

最后，在本译著出版之际，我应该首先感谢钱德拉塞卡教授允许我翻译他的这本书；其次，我应该感谢湖南科学技术出版社；再次，我还要感谢我的大哥杨建军，是他把原著买下并邮寄给我。当然，我还应该感谢我的几位研究生，正是他们的帮助使本译著能按时完稿。盐城师专物理系的王晓明讲师翻译第3、7两章，深圳长城公司的张家干翻译第6章，人民教育出版社的周国强编辑翻译第5章，其他3章和前言、附录都是由我翻译的。全书的审校、统稿工作也是由我完成的。李元杰教授对第7章作了认真的审阅，这将保证译文的科学性，我应向他表示感谢。

作者的原著肯定是一本难得的好书，但不知我们的译文能不能对得起钱德拉塞卡教授的一番深情厚意。读者的批评，将会受到译者由衷的欢迎。

[1]　F.L.Wright（1869—1959），美国著名建筑家；A.Schoenberg（1874—1951），美籍奥地利作曲家，西方现代派音乐的主要代表人物之一；M.Chagall（1887—？），美籍俄罗斯画家，他的画《乡村和我》《窗外巴黎》等蜚声欧美；John von Neumann（1903—1957），美籍匈牙利数学家、物理学家。——译者注

[2] O.Neugebauer（1899—1990），研究古代、中世纪精密科学史的著名学者。——译者注

[3] 丹麦天文学家，第谷·布拉赫的助手。——译者注

[4]　洛奇（T.Lodge，1558—1625），英国诗人及剧作家；基德（T.Kyd，558—1594），英国剧作家；皮尔（G.Peele，1558？—1597？），英国剧作家；李里（J.Lyly，1554—1606），英国作家；马洛（C.Marlowe，1564—1593）英国剧作家；格林（R.Greene，1560—1592）英国剧作家；纳什（T.Nashe，1567—1601），英国讽刺作家，剧作家及小说家。——译者注

[5]　本节所有十四行诗的译文均引用人民文学出版社出版的《莎士比亚全集》十一卷上的译文（1988），特此申明，并表示感谢。——译者注

[6] 福斯塔夫是《亨利四世》剧中一个塑造得极成功的人物。他是亨利四世皇太子的一个骑士，胆小如鼠又气壮如牛，纵情酒色又机智幽默，只要有机会就想大捞一把。但皇太子继位成了亨利五世后，他对福斯塔夫严加斥责，并把他赶出朝廷。——译者注

[7] 本·琼生（Ben Jonson，1573？—1637），英国戏剧家及诗人。——译者注

[8]　下面两段汉姆雷特的台词，借用朱生豪先生的译文（见《莎士比亚全集》9卷，第67—68页）。——译者注

[9]　埃斯奇勒斯（Aeschylus，528B.C.—456B.C.），希腊悲剧诗人；沙孚克里斯（Sophocles，495B.C.？—406B.C.？），古希腊三大悲剧作家之一；尤里皮蒂（Euripides.480B.C.？—406B.C.？），希腊悲剧作家。——译者注

[10] 黑兹里特（W.Hazlitt，1778—1803），英国散文作家。——译者注

[11] 德莱顿（J.Dryden，1631—1700），英国诗人及剧作家，1670—1688为桂冠诗人。——译者注

[12]　凯利班（Caliban），《暴风雨》中的一个人物，他是一个丑陋、野蛮而残忍的奴隶。——译者注

[13] 伍尔芙（Virginia Woolf，1892—1941），英国女作家。——译者注

[14] 埃林耿希泰脱是维也纳近郊的一个小镇，贝多芬曾在此短期住过。——译者注

[15] 参见《新约》马太福音第2章的第1和7-13节。——译者注

[16] 下面《云》的译文，引自江枫先生在《雪莱诗选》（湖南人民出版社，1987年第118页）的译文，在此向江先生表示感谢。——译者注

[17] 此处和后面《为诗辩护》的译文，参照了刘若端先生在《十九世纪英国诗人论诗》（人民文学出版社，1984年）的译文，在此向刘先生表示感谢。——译者注

[18] The Hunting of the Snark，英国作家Lewis Carroll写的诗集。——译者注

[19] 《镜里世界》（Through the Looking Glass）是英国作家卡洛尔（Lewis Carroll，1832—1898）的一本著作，书中的主角是小女孩爱丽思。——译者注

[20] 沃森于1962年获诺贝尔生理学和医学奖。——译者注

第一推动丛书: 综合系列

The Polytechnique Series

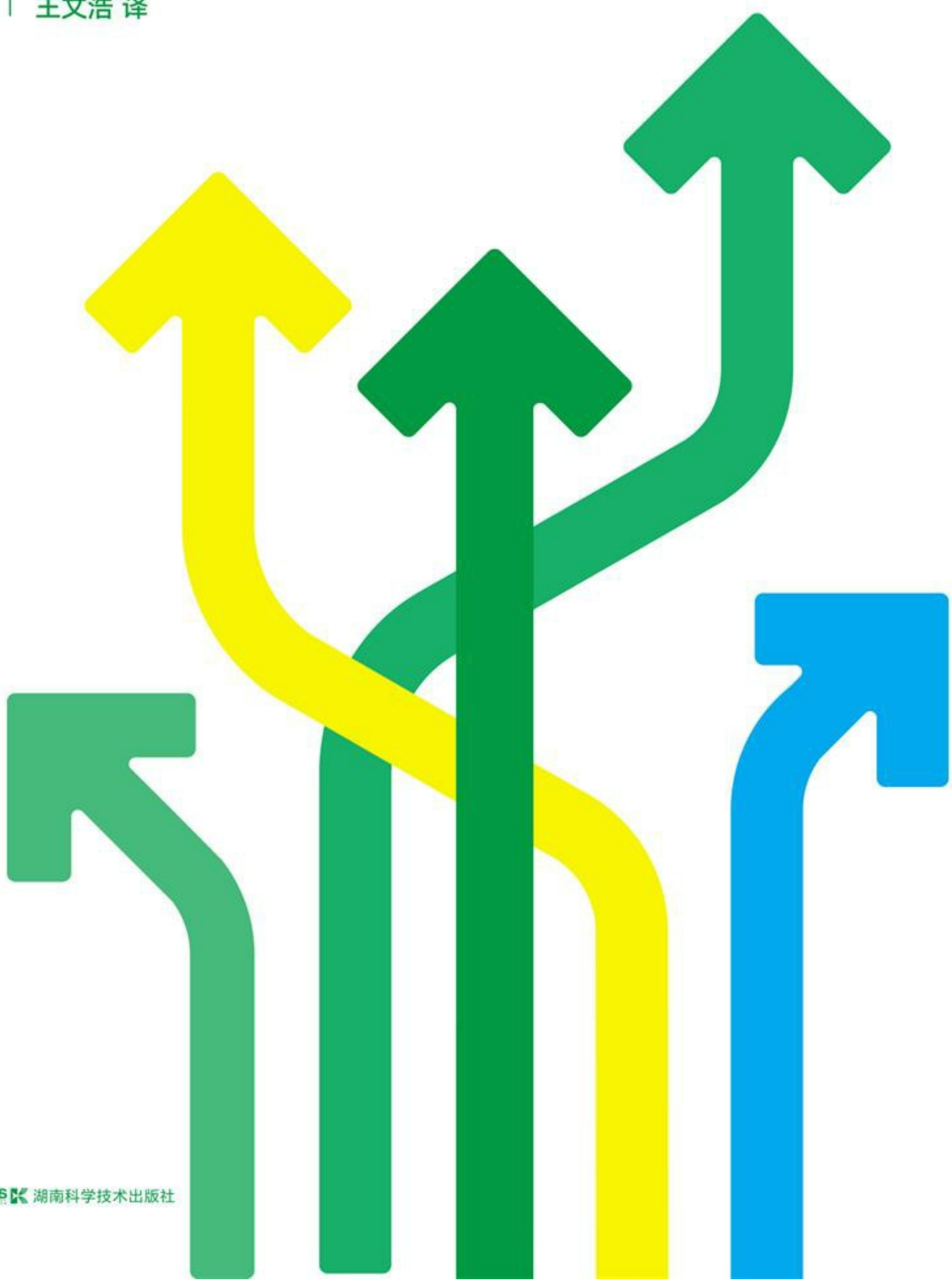
THE FIRST MOVER

下一步是什么

What is Next

[美] 麦克斯·布罗克曼 编
王文浩 译

Max Brockman



湖南科学技术出版社

版权信息

下一步是什么

编者：[美]麦克斯·布罗克曼

译者：王文浩

责任编辑：吴炜　孙桂均　杨波

装帧设计：邵年　李叶　李星霖　赵宛青

ISBN：9787535794376

湖南科学技术出版社
HUNAN SCIENCE & TECHNOLOGY PRESS

目　录

版权信息
总序
再版序 一个坠落苹果的两面：极端智慧与极致想象
序言
第1章 我们将撤往北缘地区？
第2章 镜像神经元：我们天生就合乎道德吗？
第3章 如何增强人自身
第4章 我们在非自然宇宙中的位置
第5章 暗能量到底是什么？
第6章 青少年中社会脑的发展
第7章 观察大脑的互动
第8章 是什么使得大观念如此牢固？
第9章 道德心灵的生理学基础
第10章 语言是如何影响我们思考的？
第11章 记忆增强，记忆擦除：我们过去的未来
第12章 想象的极端重要性
第13章 脑时间
第14章　超越大脑：为什么智人要从树上下来，为什么其他动物不跟着学？
第15章 我们当中的外来者
第16章 社群性昆虫是如何变得具有社会性的？
第17章 人类的灭绝和进化
第18章 为什么专业化没有导致科学四分五裂？
致谢

总序

《第一推动丛书》编委会

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难的进步。这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知做了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人

的强国梦。然而可以说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是，毋庸讳言，在一定的范围内或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的结果的接受和承认，而不是对科学的原动力——科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上不隶属于服务于神学，不隶属于服务于儒学，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化和地域差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学精神和科学思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步做一点推动。丛书定名为“第一推动”，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现自身如何成为自身的主宰。

再版序
一个坠落苹果的两面：极端智慧与极致想象

龚曙光

2017年9月8日凌晨于抱朴庐

连我们自己也很惊讶，《第一推动丛书》已经出了25年。

或许，因为全神贯注于每一本书的编辑和出版细节，反倒忽视了这套丛书的出版历程，忽视了自己头上的黑发渐染霜雪，忽视了团队编辑的老退新替，忽视好些早年的读者已经成长为多个领域的栋梁。

对于一套丛书的出版而言，25年的确是一段不短的历程；对于科学研究的进程而言，四分之一个世纪更是一部跨越式的历史。古人“洞中方七日，世上已千秋”的时间感，用来形容人类科学探求的速律，倒也恰当和准确。回头看看我们逐年出版的这些科普著作，许多当年的假设已经被证实，也有一些结论被证伪；许多当年的理论已经被孵化，也有一些发明被淘汰……

无论这些著作阐释的学科和学说属于以上所说的哪种状况，都本质地呈现了科学探索的旨趣与真相：科学永远是一个求真的过程，所谓的真理，都只是这一过程中的阶段性成果。论证被想象讪笑，结论被假设挑衅，人类以其最优越的物种秉赋——智慧，让锐利无比的理性之刃，和绚烂无比的想象之花相克相生，相否相成。在形形色色的生活中，似乎没有哪一个领域如同科学探索一样，既是一次次伟大的理性历险，又是一次次极致的感性审美。科学家们穷其毕生所奉献的，不仅仅是我们无法发现的科学结论，还是我们无法展开的绚丽想象。在我们难以感知的极小与极大世界中，没有他们记历这些伟大历险和极致审美的科普著作，我们不但永远无法洞悉我们赖以生存世界的各种奥秘，无法领略我们难以抵达世界的各种美丽，更无法认知人类在找到真理和遭遇美景时

的心路历程。在这个意义上，科普是人类极端智慧和极致审美的结晶，是物种独有的精神文本，是人类任何其他创造——神学、哲学、文学和艺术无法替代的文明载体。

在神学家给出“我是谁”的结论后，整个人类，不仅仅是科学家，包括庸常生活中的我们，都企图突破宗教教义的铁窗，自由探求世界的本质。于是，时间、物质和本源，成为了人类共同的终极探寻之地，成为了人类突破慵懒、挣脱琐碎、拒绝因袭的历险之旅。这一旅程中，引领着我们艰难而快乐前行的，是那一代又一代最伟大的科学家。他们是极端的智者和极致的幻想家，是真理的先知和审美的天使。

我曾有幸采访《时间简史》的作者史蒂芬·霍金，他痛苦地斜躺在轮椅上，用特制的语音器和我交谈。聆听着由他按击出的极其单调的金属般的音符，我确信，那个只留下萎缩的躯干和游丝一般生命气息的智者就是先知，就是上帝遣派给人类的孤独使者。倘若不是亲眼所见，你根本无法相信，那些深奥到极致而又浅白到极致，简练到极致而又美丽到极致的天书，竟是他蜷缩在轮椅上，用唯一能够动弹的手指，一个语音一个语音按击出来的。如果不是为了引导人类，你想象不出他人生此行还能有其他的目的。

无怪《时间简史》如此畅销！自出版始，每年都在中文图书的畅销榜上。其实何止《时间简史》，霍金的其他著作，《第一推动丛书》所遴选的其他作者著作，25年来都在热销。据此我们相信，这些著作不仅属于某一代人，甚至不仅属于20世纪。只要人类仍在为时间、物质乃至本源的命题所困扰，只要人类仍在为求真与审美的本能所驱动，丛书中的著作，便是永不过时的启蒙读本，永不熄灭的引领之光。虽然著作中的某些假说会被否定，某些理论会被超越，但科学家们探求真理的精神，思考宇宙的智慧，感悟时空的审美，必将与日月同辉，成为人类进化中永不腐朽的历史界碑。

因而在25年这一时间节点上，我们合集再版这套丛书，便不只是为了纪念出版行为本身，更多的则是为了彰显这些著作的不朽，为了向新的时代和新的读者告白：21世纪不仅需要科学的功利，而且需要科学的审美。

当然，我们深知，并非所有的发现都为人类带来福祉，并非所有的

创造都为世界带来安宁。在科学仍在为政治集团和经济集团所利用，甚至垄断的时代，初衷与结果悖反、无辜与有罪并存的科学公案屡见不鲜。对于科学可能带来的负能量，只能由了解科技的公民用群体的意愿抑制和抵消：选择推进人类进化的科学方向，选择造福人类生存的科学发现，是每个现代公民对自己，也是对物种应当肩负的一份责任、应该表达的一种诉求！在这一理解上，我们将科普阅读不仅视为一种个人爱好，而且视为一种公共使命！

牛顿站在苹果树下，在苹果坠落的那一刹那，他的顿悟一定不只包含了对于地心引力的推断，而且包含了对于苹果与地球、地球与行星、行星与未知宇宙奇妙关系的想象。我相信，那不仅仅是一次枯燥之极的理性推演，而且是一次瑰丽之极的感性审美……

如果说，求真与审美，是这套丛书难以评估的价值，那么，极端的智慧与极致的想象，则是这套丛书无法穷尽的魅力！

献给我的父母

序言

麦克斯·布罗克曼

纽约

2009年1月

我们大多数人从长期的经验中学到，与下一代进行思想上的接触，深入了解我们当今时代的问题以及未来社会将要面对的问题，这在社会各个领域都是非常重要的。在科学领域，这项工作尤为有价值，因为很多重要的发现都是该领域初出茅庐的年轻人做出的。编辑这本简明文集的初衷就是想一睹当今这些耀眼的青年科学家在做什么、想什么。

这里专载的18位青年科学家所从事的研究各不相同，但所涉问题将会对我们的生活——甚至对我们如何看待自己以及我们如何看待我们在宇宙中的位置——产生长期和深远的影响。他们的思想最终将有助于重新确立我们是谁，我们是什么。

为了形成这份名单，我走访了当今一些领袖级科学家，请他们提供各自学科领域的新星：这些人在各自的研究中正处于解决和提出科学上一些最棘手问题的前沿。这份名单可以说是未来一代科学家名录的精华版。

我要求每一位入选者就其自身的研究领域撰写所面临的问题。他们的文章特别新颖，因为这些作者里多数人很少有时间或机会为广大的非专业读者写东西。这个名单里有：

•戴维•伊戈尔曼，贝勒医学院知觉与行为实验室主任，从事大脑如何感知时间的分析；

•卡捷琳娜•哈尔瓦蒂，马克斯•普朗克进化人类学研究所的古人

类学家，研究原始人类在过去遭到灭绝的证据，以及它对我们自己这个物种在未来的意义；

• 马修•利伯曼，加州大学洛杉矶分校的心理学副教授，研究是否是我们的大脑的生理结构使得我们更有可能形成并坚持某些思想；

• 肖恩•卡罗尔，加州理工学院物理学方面的高级研究助理，讨论为什么我们仍不知道我们这个宇宙的起源以及时间的方向；

• 劳伦斯•史密斯，加州大学洛杉矶分校地理系教授兼系副主任，地球与空间科学系教授，他认为气候变化可能导致陆地上迅速形成新的经济上活跃的北部地缘带；

• 莱拉•博罗迪茨基，斯坦福大学心理学、神经科学和符号系统系助理教授，研究我们的语言如何影响我们的思维方式；

• 萨姆•库克，麻省理工学院神经科学博士后助教，研究我们何时以及如何能够操纵自己的记忆——以及我们是否应该这么做。

这些年轻研究人员的工作热情以及对科学的热爱是显而易见的。他们对新思想的大胆探索，对现有知识边界的突破性尝试令人振奋。我希望《下一步是什么》这本书会给读者在思考面对未来我们该做些什么样的准备方面提供一个指导性的开端。

第1章
我们将撤往北缘地区？

◎劳伦斯·史密斯

劳伦斯·史密斯（Laurence C.Smith）

1996年在康奈尔大学获得地球与大气科学博士学位，目前是加州大学洛杉矶分校地理系教授兼系副主任，地球与空间科学系教授。他已在包括《科学》和《自然》等杂志上发表50多篇研究论文。2006年，他向美国国会介绍了北半球气候变化的可能影响；2007年，他的工作出现在联合国政府间气候变化问题研究小组（IPCC）第四次评估报告的显要位置。

史密斯关于北半球气候变化的工作得到古根海姆基金会、美国国家科学基金会和美国航空航天局的资助。他获得的荣誉包括美国航空航天局青年科学家奖（2000年），美国航空航天局总统奖（2002年），并从洛克菲勒基金会得到享用Bellagio官邸的荣誉（2007）。

正如许多其他的文化变迁——经过长期积累，然后迅速爆发——那样，人类——包括美国的大多数人——最终承认全球变暖是真实的。

改变公众的观点并非易事。它是数以千计的科学家辛苦工作了30多年积累起来的成果。这些成果由联合国政府间气候变化问题研究小组（IPCC）以大型综合报告形式，分别于1990年、1995年、2001年和2007年逐步传达给全世界，它们展示了科学上前所未有的组织水平。这些报告以无可辩驳的证据显示，一种新的人为的气候现在已呈压倒性优势。

对民意改变起关键作用的是一群热心的“第三文化”科学家——其中包括美国航空航天局戈达德空间研究所的詹姆斯·汉森、俄亥俄州立大学的朗尼·汤普森、宾夕法尼亚州立大学的理查德·阿利和科罗拉多大学

的马克·塞雷兹——他们天才地抓住了最重要的发现，并通过图书、访谈、YouTube网站和《滚石》等大众杂志将其传递给公众。这些在扩大服务于公众方面的努力代表了科学文化的重大转变。在20世纪90年代中期我还是研究生的时候，我曾亲眼看到著名天文学家兼作家卡尔·萨根受到的、来自于他的同事对他在宣传科学工作方面的努力所表现出的广泛而又微妙的轻蔑。但是今天不同了，特别是在气候变化科学方面，向公众宣传已成为科学界同仁工作的一部分，并且受到同事的赞赏和效仿。

其他一些事件，主要是那些不可预见的事件，在转变公众观念方面也具有突出的作用。卡特里娜飓风带来的恐怖情形——不管这次飓风的成因如何——已通过电视和电脑屏幕让数以百万计的国民感到不安。戈尔在2000年竞选总统的失败使他能够抽身出来全力投身于环境保护事业，并在2006年投资拍摄了电影《难以忽视的真相》。为此，他与IPCC一起赢得了2007年度诺贝尔和平奖。2006年，沃尔玛做出的采用并大力推广绿色经营技术的决定已经得到数百万人的积极响应，其中很多人都是受到了戈尔电影的启发。在我的老家加州，共和党州长施瓦辛格宣称："气候争论已经结束"——不论是从科学的角度还是民意的角度。他说对了。

对于过去，我们可以列举出各种证据来证明环境形势的严峻性，但对于未来，这些证据还是那么具有说服力吗？争论（如果算的话）非常尖锐，但战线已推向敌方。像"这是真的吗？"和"是不是我们错了？"这类问题现在已经被替代为"会发生什么？""发生在哪里？""会有多快？"和"我们该做些什么？"等确定性问题。科学将我们带到了这些问题前，但我们的答案将远远超出科学的范畴。说它关系到21世纪人类生存的全球格局绝不是危言耸听。

因此，到底会发生什么？以下是我们目前所知道的：第一，气候变暖才刚刚开始。我们有90%的把握肯定：如果以目前或超过目前的速度继续排放温室气体，那么21世纪气候变化之剧烈将远远超过我们曾经历过的任何一个世纪。[1] 考虑到人口增长因素或到下个世纪的温室气体排放，按物理学基本原理可推知，地球的气候必将继续升温，到21世纪末全球将平均升高1.8℃～4.0℃（3.2～7.0），除非有些至今没有发现的非线性气候因素在起作用。[2] 气温到底会上升多高取决于我们向大气排放多少二氧化碳。这里，较低的值是IPCC的乐观估计，即假设全球

人口稳定并且采用清洁能源技术。较高的值是基于对化石燃料的依赖有增无减的估计。

这些温度变化听上去似乎并不大，但实际情形并非如此。即使是上述最乐观的估计（温升1.8℃），也是我们在20世纪经历的3倍。此外，由于温室气体在大气中的长寿命以及世界各大洋对其反应的迟缓，我们已经“锁定”了这种变暖的大部分事实，就是说，无论我们制定的政策有什么样的变化，全球气温上升将持续到2030年已是不争的事实。即使我们能将温室气体排放量限定在2000年的水平，这种气候变暖依然会延续到21世纪中叶。但从长期来看，政策变化将产生巨大的影响：预计到2100年，温度上升将只有目前锁定值的20%。从这一点上看，通过积极的社会行动来延缓气候变暖还是有可能的，尽管我们不能阻止它。

较高的气温将加剧水分蒸发，使土壤变得干燥，并使旱灾频发，尤其是在南北纬度20°～40°之间的两大宽阔带，情形就更是如此。气候变暖将使美国西南部、欧洲南部和东部、非洲南部和南美的东部等地的极端干旱天数大大增加。[3] 空气中水汽也将增加，按照克劳修斯—克拉珀龙方程，气温每提高1℃，大气的含水量上升7%。由于天气系统中水蒸气的增加，极端降水事件——即洪水——的频率将随之上升。热浪造成的致命的大面积停电将经常发生，就像2003年在法国，2006年在美国，2007年在日本一再出现的情形，热浪造成死亡率攀升。海平面持续抬高（它的增长现在是每年大约3mm），唯一不确定的仅仅是它上升得有多快、有多高。低海拔的沿海地区，包括佛罗里达州、荷兰、岛屿国家和贫困的孟加拉国，将面临在今后几十年内被淹没的危险。

如果你看过影片《难以忽视的真相》或在报纸上读到过气候变化的故事，你一定已经了解大多数这类坏消息。关于21世纪气候变化的这些报道，除了有关飓风和森林火灾的属于推测性的之外，大部分都属于科学预言在先。然而，即使这些都还不是我们的气候模型做出的最严厉预警。气候最强有力的变化将席卷整个北部高纬度地区，北纬45°线基本上穿越美国北部、加拿大、俄罗斯和欧洲。在这条线以北，气候变化将是地球上任何其他地区无法相比的：气温将上升全球平均水平的近一倍（主要由暖冬造成），降水也将大幅增加。

在极北部地区，这些影响已显露端倪：北极海冰融化，北极熊溺亡，因纽特猎人失业。所有这些都是全球变暖的象征。北极地区气候变

暖的速度和严重性确实非常值得注意。当然，北极地区相对较小且人口稀少，其原生态的社会和经济结构对我们生存的其他地区的影响永远属于边缘性的。但其南部区域，即“北部边缘”（northern rim，以下简称“北缘”）附近的广大地区及邻近洋域——美国、加拿大、丹麦、冰岛、瑞典、挪威、芬兰和俄罗斯境内区域——问题就要严重得多。如同北极的情形，这些区域的气候变化已经开始出现。在21世纪里，这一地带——占近30%的地球陆地面积，现存的最大森林区域，也是最大的未开发矿产、水资源和能源蕴藏基地，有近1亿人口——将经历最深刻的生态的和社会结构的变迁。

这些北纬地区对南方居民从来就不具有吸引力，这有几个原因。首先是日照时间具有很强的季节性。在极北地区和内陆深处，永久冻土层使建设成本难以承受，土壤常年渍涝，使潮湿的土地成为数以亿计的蚊子的天堂。作物生长季节短，农业产量低，大部分土地属于山地类型。对于适于南方生长的生命形式——植物、动物和人类来说，唯一最大的障碍就是令人脑袋发木的隆冬的严寒。这里的夏天是温暖的，甚至酷热（好在整个北缘附近都能用空调）；但冬天却是严霜冰冻。落叶乔木干裂死亡，蟾蜍在泥床上冻得像土坷垃，在-40℃（在这种低温下，摄氏和华氏温标已趋同）的严寒下，压缩机失灵，钢铁断裂，人工建设已不可能。[4]

经历过“-40℃”的人都对这种恶劣环境感到惧怕和痛恨，它使人类的活动被迫停止。所有来自北缘的人——白马镇的老板、艾伯塔省的克里族渔民、俄罗斯的卡车司机和赫尔辛基的退休老人——均向我描述了这种感觉。他们对近期气候变暖带来的各种问题和机会的态度颇为复杂，但有一点是共同的：那就是都相信-40℃的天气正变得越来越少。

全球变暖使得北缘的隆冬严寒趋缓，那是不是意味着这里可能是人类社会新的栖息地呢？这个想法并非那么荒诞：据2007年8月8日《华尔街日报》的一篇文章介绍，在纽芬兰和拉布拉多，投机性房地产交易在2007年大幅上涨。同样，据2007年12月1日的《金融时报》报道，挪威北部、瑞典和俄罗斯的房地产市场亦呈上扬趋势。但在你急着点开Realtor.com网挑选安克雷奇或温尼伯附近地区的不动产之前，先听好了：是的，房地产交易是在增长，但不是处处如此。就如同几千年来人类的扩张一样，它的方向取决于我们作出的选择和此前留下的历史和地理的印记。

我的加州大学洛杉矶分校的同事贾里德·戴蒙德（Jared Diamond）在他的《崩溃》一书中，通过对人类历史的梳理，确认了导致现存社会有可能走向衰亡的5个主要因素：环境破坏、贸易伙伴的失去、以邻为壑、气候变化和社会对环境问题做出的选择。所有这些因素，单独或组合，都可能引发社会的崩溃。就这个问题而言，是什么能够使新的社会结构有可能成功地建立起来？首先是经济机会，接下来依次是环境适宜性，投资和贸易的机会（它隐含着军事安全和完善的法治，没有这些，投资者就会畏葸不前，贸易就不会稳定），友好的邻邦和乐意定居者。

目前，在北缘地区，这些要求仅在不同程度上得到部分满足。充足的经济机会要求自然资源——化石燃料、矿产、鱼类和木材——以商品形式存在。事实也确实如此，目前北缘地区的国内生产总值主要来自这些资源的开发，其次是政府服务的贡献。这里邻国之间通常是友好的，相对于世界其他地区，所有8个北缘国家的内乱程度较低，且边境态势友好——虽然芬兰对与俄罗斯之间的漫长边境时有烦恼，俄罗斯则担心其人口稀少的东翼受到美国和（尤其是）中国的威胁。但不管怎么说，自第二次世界大战以来，这8个国家之间不存在严重的军事入侵。他们中的7个（俄罗斯是个例外）享有世界上最稳定的政治制度和法治环境。

接下来要考虑的是环境适宜性、贸易和人口等因素。气候变暖很可能使人类向北缘扩张的最大的环境制约因素——严冬——变得可以忍受。气候变暖还将缓解其他一些问题，譬如较短的生长季节；但也会带来另一些问题，譬如病虫害。但与暖冬的影响相比，这些都是次要的。因此，气候变化，这个使过去社会崩溃的五大关键因素之一，在北纬地区实际上起着孕育新社会的作用。

至于贸易和人口，也都好办，这些因素主要取决于市场、基础设施和人口发展趋势。虽然商品价格存在着波动，但从长期（例如21世纪和下世纪）趋势看，我们可以放心地说，世界对淡水、矿产、能源、食品和木材的需求将持续走高。但需求本身并不一定带来贸易，还必须有基础设施，没有这些商品将无法进入市场。人口因素同样如此，没有一定量的人口就不存在劳动力市场。而要满足充足的人力资源条件，则要求当地人口呈增长态势，或通过移民，或两者兼而有之。下面我想通过对今天北缘地区基础设施和人口发展趋势方面的强烈对比，来说明我对改变人类向北缘扩张的地域模式的一种希冀。否则的话，我们就会在某种

程度上打乱原有的体系——20世纪我们曾两度如此，而且这种干扰首先造成的就是地域环境上的反差。

北缘地带长期以来资源丰富但人口稀少，这对各国中央政府是一种不可抗拒的诱惑。多年来，各国政府在改善基础设施、促进人口和经济增长方面的努力一直受到不同意识形态的左右，结果可谓好坏参半。此外，各国政府对待原住民的方式也有巨大差异。近年来，美国和加拿大在这方面做得最好，其次是斯堪的纳维亚，最差的是俄罗斯。但是与20世纪人类做出的两次重大选择比起来，所有这一切都显得相形见绌。那两次选择彻底改变了人类在北缘的存在：一次是第二次世界大战期间美国军队决定占领加拿大，另一次是1929～1953年，斯大林决定在横跨西伯利亚地区建立劳改营和流放地。

但作为一种强制移民措施，这一计划取得了巨大成功。到20世纪50年代初，劳改营人口达到250万，其中大部分是政治犯或轻刑犯。[5] 他们下矿井，伐木，修建公路、铁路和工厂。即使熬到刑满释放，出狱者也不允许回原籍。城镇发展得很快，到80年代末，这些城镇已发展为大城市，它们横跨地球上最寒冷的地域：新西伯利亚、鄂木斯克、叶卡捷琳堡、哈巴罗夫斯克、车里雅宾斯克、克拉斯诺亚尔斯克、诺里尔斯克、沃尔库塔市。俄罗斯已使西伯利亚城市化。

今天，这些城市的前途未卜。它们的位置选定是随意的，而不是出于经济可行性的务实要求。它们所在的地点让人不可思议：环境恶劣，彼此间距离遥远，贸易伙伴远不可及，基础结构延绵千里，供应链岌岌可危，需要从莫斯科得到巨额补贴才能运转。苏联担负社会主义建设的规划者们就是这样，通过在不毛之地建立起这些城市而使国家经济体系背上了沉重负担。菲奥娜·黑尔和克利福德·杰迪在他们的《西伯利亚诅咒》（Fiona Hill and Clifford Gaddy，*TheSiberian Curse* ）一书中解释说，“寒冷成本”压弯了苏联经济，并最终促使苏联在1991年解体。

苏联解体后，对这些城市的补贴也随之消失。在整个90年代，西伯利亚巨大城市群的人口流失速度比底特律不景气的裁员年的人员流失还快。今天，人口已有回稳的迹象，有限的繁荣主要来源于高油价。在21世纪里，第二次针对西伯利亚大型基础设施建设的更为明智的尝试可能很快就要出现，因为俄罗斯和中国都已看上了它那巨大的自然资源。目前已有初步迹象：普京正式开放了6200英里长的莫斯科到符拉迪沃斯托

克的高速公路，这是世界上最长的公路。俄罗斯军事学者已提出在西伯利亚远东地区实行“自由经济区”概念，开放该地区巨大的木材资源以便利用中国的资金进行发展。但是，那里的人口在继续减少，俄罗斯族人和原住民都是如此——高死亡率、自杀和极度贫困导致人口增长率持续走低。西伯利亚任何新的人口扩张都需要大量的基础设施建设，需要更明智的人口发展计划，对失败的城镇选择放弃或迁移，只有这样才能使今天的人口负增长的局面逆转。

在北美则是另一番景象。最初少数白人殖民者与土著群体共同生活，他们通过开矿、捕猎、打渔和小农场经营来不断开拓生存空间。第二次世界大战改变了这一切。从战时需要出发，美军在北缘地区修建了大型基础设施——机场、道路、基地、管道、港口、雷达站以及遍布阿拉斯加、加拿大、格陵兰和冰岛的所有城镇。仅在阿拉斯加就有6万美军和承包商。在加拿大西北部，友好地驻扎着总共约4万士兵和平民工作人员——他们为这一人口稀少地区输入了大量人口。加拿大政府从渥太华看着这一切：美国用西北分期路线（Northwest Staging Route）、阿拉斯加公路、卡诺尔管道和其他几十项工程项目改变着这一地区。[6]美军工程的印记至今仍影响着北缘居民的生活模式和经济活动。

土著群体则完全被边缘化，看着他们的家园不断萎缩。在加拿大，他们曾忍受着文化遭到破坏的重新安置计划。但20世纪70年代初以来，他们不断通过法律行动（要求土地所有权）来恢复其领地管辖权和矿业管理权，现已基本成功。美国和加拿大的原住民正在形成蓝色西装商业机构，并不断加强其经济和政治影响力，他们在北缘地区拥有最高的人口增长率。如果美国持续干旱下去，我和我的气候变化难民同胞将无法指望在北方边境立足，因为所有好的地点都已瓜分完毕。

那么，在向北扩展这一点上，上述对比给了我们什么启示呢？美国和加拿大有合理的基础设施、有效的法律法规以及快速增长的安居乐业的家庭人口，这两个国家具有很好的扩张潜力。俄罗斯则不然，它的基础设施严重滞后，人口严重下降。斯堪的纳维亚国家则以其发达的公路网、港口和大学教育，已准备好从冬季气候变暖的过程中获益。

为了免得你误解这篇文章，让我把一些事情讲清楚：我所描述的转变是从不适宜居住的土地迁徙到适于居住的土地。这决不是建立北部的乌托邦。如果我们可以待在现在生活的地方，享受从北纬50°到南纬45°

之间更广大、更友好的土地，那么国际社会将远远和平得多。我们这个星球的自转轴倾斜23.5°，这决定了总会存在一些黑暗和寒冷的高纬度地区，即使温室效应使得马尼托巴省丘吉尔地区的2月热得（比方说）赛过明尼阿波利斯的2月。事情远没到木已成舟的地步：是的，我们是受到气候明显变暖的影响，但还不至于（至少目前）都必须迁居到雅库茨克。我们的选择是：如何改造这一地区需要通过明智的、深谋远虑的计划来决定，这一计划应是基于IPCC对温室气体排放的乐观或悲观的估计。往好了说，北缘地区将变得可以忍受，而不是天堂。我不建议你在拉布拉多置地。但在密歇根州也许可以考虑。

第2章
镜像神经元：我们天生就合乎道德吗？

◎克里斯蒂安·凯泽斯

克里斯蒂安·凯泽斯（Christian Keysers）

1973年出生于比利时，曾在德国和波士顿学习心理学和生物学。2000年获苏格兰圣安德鲁斯大学神经系统方面博士学位，随后去意大利帕尔马工作，在那里他对听觉镜像神经元的发现做出了贡献，并将镜像神经元的概念推广应用到情绪和感觉方面。他是荷兰格罗宁根大学医学中心神经影像中心主任和社会脑研究方向的全职教授。他曾获得欧洲委员会颁发的居里夫人卓越成就奖，是《社会神经科学》杂志的副主编。

共享回路：你如何侵入我的大脑

在20世纪90年代，帕尔马大学的意大利神经科学家贾科莫·里佐拉蒂（Giacomo Rizzolatti）、维托里奥·加莱塞（Vittorio Gallese）、莱昂纳多·弗加西（Leonardo Fogassi）及其同事在研究大脑如何控制我们的行动这一课题时有一项重要发现，他们使用非常薄的电极来测量猴子大脑中运动前皮质区域单个神经元的活动。当猴子抓握或操作物体时，这个区域的神经元就活跃起来。例如，当猴子抓住一颗花生时，某些神经很活跃；当猴子剥开花生壳时，则另一些神经元变得活跃。正是这些神经元的活动触发了猴子的动作。人类也有运动前皮质，如果一名外科医生在进行病人手术时刺激了这个区域，病人就会报告说有想要执行某些操作的感觉。运动前皮质是我们自愿行动和控制我们自己身体——我们的自由意志堡垒——的关键。

当实验中一只猴子抓起一颗花生交给另一只猴子时，奇迹出现了：对抓住花生起反应的神经元在猴子只是看着别人执行相同操作时也会出

现同样的反应。起初，佐拉蒂等人不相信他们的这一发现。大脑中涉及自愿行动的区域怎么可能会在观看其他人的行为时产生反应？最终让他们信服的是这种反应的显著的一致性：对执行特定操作（如抓物）起反应的神经元会对看到的特定行动产生反应。这不可能是巧合。

实验者后来证明，即使是其他人动作发出的声音也会激活与这些特定动作相联系的神经元。事情变得越来越明显：猴子的大脑可以将实验者的动作转变成猴子将用来执行同样动作的引擎程序。

我们称这些神经元为镜像神经元，因为通过它们，猴子脑部的引擎活动映射了他人的行为。

一系列实验表明，人类也有类似的神经系统。当我们看到了其他人、动物甚至机器人的行为时，我们自身行为的引擎表达就被激活。

如果我们看到了其他个体的行为时我们自身行为的引擎表达会被激活，那么我们为什么不总是公开地模仿别人？答案是，当我们正在观察他人的行为时，神经门似乎锁住了我们的引擎区域的出口，不让身体模仿看到的动作。但在这扇门的背后，我们的大脑可以隐蔽地分享我们周围的人的行动。我们不再只看到别人的动作，我们也感到我们内心的动作，就像我们正在做着同样的事情。例如，如果我们处在比赛的起跑位置上，这时旁边对手的假动作就很可能自动触发我们自身的起跑。当我们看到别人跳舞，我们常常会情不自禁地感到自己的身体也在律动。我们的引擎系统就像是我们的自由意志的活动中心，每当我看到你的动作，你便渗透到这个活动中心。你的行为成了我的；我的行为成了你的。

这种现象不只是限于身体运动。譬如说，当有人拍了下我的肩膀，我的感觉皮质让我感受到一种感觉。但有意思的是，即使我看到有人拍了另一个人一下，我的大脑的同一区域也会被激活。如果我割破了自己的手指，我的扣带皮质和前脑岛将记录下这种痛苦。但如果我看到你割破了你的手指，我的大脑的这些部位也会活跃起来。虽然这种间接感受不像我们体验自己的感觉那样强烈，但我们毕竟隐隐地感觉到其他人的感受。当我看到你割破你的手指时，我感觉到我的手指也隐隐作痛。当我看到詹姆士·邦德的肩膀上爬着一只大蜘蛛时，我自己的肩膀也会痒痒。

不仅如此，我们的情绪似乎也遵从类似的法则。例如，当我闻到难闻的气味时，我的岛叶产生厌恶感。但当我看到你的脸上表现出厌恶的神色时，我的岛叶区域同样会活跃起来，就好像我体验到你的厌恶。这些共同的神经活动与我们的个人主观经验是密切相连的：当我们听到有人在笑时，我们的情绪也会好起来；当你的朋友在哭时，你也会感到悲伤。他人的情绪是有传染力的，因为我们在看到他人的情绪表现时我们的大脑会激活我们自己的情绪。

这些大脑回路可以防止我们在看到其他人的境遇时表现得像个“局外人”。事实上，我们可以在内心感受到他们的动作、感觉和情绪，就好像我们就是他们。他们已经成为我们。

共享回路允许我们向他人学习

这些共享回路有什么好处呢？人类通常将成功归因于自身具有合作和相互学习的能力。狩猎就是一个很好的例子：有了长矛和协调能力，人类可以降伏水牛或猛犸象。虽然媒体经常美化个人的天才，诺贝尔奖也只是奖给新思想的发现者，但大多数有用的东西（如长矛）是几千年中技术缓慢进步的结果——从较有经验的老师那里学来，然后加入自己的创新，再教会下一代。

不知怎的，我们的大脑就是会向他人学习。这个过程远不是那么不起眼。例如，为了学会制作长矛，我们必须将对他人动作之所见变成很不相同的某种东西——以类似方式来移动我们自己的手所需的神经脉冲。镜像神经元似乎已经解决了这个艰巨的任务：每次我们看到一个动作，镜像神经元就会将所见转换为重复这一动作所需的引擎命令。当我们看到行家一个动作接一个动作地进行着一连串的操作，直到最后制作出了长矛，我们的大脑会以相同的顺序激活类似的行动，将捡起石块，磨尖它，捡起木棒，将石块绑在木棒上等熟悉的动作组成一整套新颖的长矛制作程序。

与别人的感情交流是这一过程的重要因素。几乎所有动物都是基于反复尝试—纠错来学习的。当我们共享他人的行动和情感时，这种古老的机制就成为一种社会性的学习。如果我们看到某个人品尝一种不熟悉的水果，并流露出高兴的神色，我们的大脑就会共享这个动作，并产生积极的后果——好像我们自己吃了这种水果并喜欢上它；如果我们看到

某人吃了之后表露出厌恶的神情，我们也会分享这种负面体验。我们的替代性的试错学习机制会告诉我们，吃了这种水果是好还是不好，这为我们提供了一种没有中毒风险的学习的好处。

由此，这些大脑回路模糊了你我经验上的明确界限。我们的经验汇集到共同的知识海洋里，我们称之为文化。随着语言、书籍和电视的出现，这种共享成为全球性的，它使我们能够跨越时间和空间来交流经验。

共享回路创建道德本能

我们更多时候是被动的观察者，我们的许多决定影响到他人。想象一下，我和另一个饿汉在一起，我发现了一块食物。我是一个人独吞了它，还是和他一起分享？

从个人主义的角度来看，独吞似乎是最合理的决定。它使我免受饥饿，唯一不足的是另一个人会饿死——但这不是我的问题，不是吗？如果我与他分享食物，我可能希望日后其他人也会觉得有义务回馈于我——但谁知道，他也可能不会。

不过，由于存在这些共享回路，这种情形是不会出现的。因为不论是我们感觉到自己的疼痛还是目睹他人的疼痛，我们大脑的同一区域都会变得活跃，这意味着替代性地分享他人的感受不是一种抽象的概念，而是我们自己的低声部共鸣。如果我吃独食，我就不仅会看到而且会体验到我同伴的痛苦，而如果我与同伴分享食物，我就会分享他的喜悦和感谢。因此，我做决定不再只是由我的饥饿来引导，而是由我的同伴的痛苦和快乐带给我的真实的痛苦和快乐体验来引导。

更重要的是，虽然这些大脑机制的力度因人而异，也因情形而异，但所有个体都会表现出某种程度的共享。因此，我们有理由认为，当其他人在决定是否让我分享食物时，也同样会考虑到我的痛苦。事实上，对社会各个方面的实验结果表明，在世界大多数地方，人们往往倾向于与其他人共享，而不是将一切财富留给自己。

共享回路使我们能够与他人分享知识，因此它对于认识人的本质还有另一层深刻含义：它们奠定了直觉利他主义的基础。大多数文化里都

有所谓道德黄金法则。例如，基督教要求“所以无论何事，你们愿意人怎样待你们，你们也要怎样待人。因为这就是律法和先知的道理”（马太福音7：12）。伊斯兰教也认为，“你们任何人都不算（真正）归信，直到他为他的兄弟祈望，如同他为自己所祈望的一样”（穆罕默德，脑威圣训第13条）。我认为，使我们分享他人的痛苦和喜悦的大脑机制是我们很容易直觉地信奉这一格言的神经生理基础。我们的大脑与生俱来就是合乎道德的。

这并不是说，人类不会互相伤害。事实上，如果我们的个人利益与他人的利益直接冲突，我们的趋利愿望可能就会超越同情心。不幸的是，人类的智慧已经设计出各种方法以便在那些不能达到我们目的的情形下减弱我们的同情心。在军事上，将军与士兵之间的距离使前者对后者遭受敌方伤害带来的痛苦变得漠然，并能够下达于己有利的命令。同时，层层下达的指挥链将直接见证痛苦的士兵身上的道义责任脱去了。军事上的高效性就这样绕过了同情心。远距离杀伤性武器的发展也有类似的作用。看清同情心的生物学本质有助于我们认识到这种距离带来的风险和问题，并为我们如何将同情心的自然机制融入我们的制度建设指明了方向。

人类是进化的结果，进化总是偏向那些能够留下更多后代的个体，而不是那些在服务于他人时忘记自身利益的利他主义个体。乍一看，驱使我们分享他人情绪的大脑似乎与适者生存法则相悖。但人类是群居动物。想象一下，一种是家庭成员可以共享其他成员的行动、感觉和感情的家庭，另一种是不具有这种能力的家庭。在后一种家庭中，兄弟间彼此偷窃，尔虞我诈，家庭成员之间很少能够相互学习。而高度共享的家庭则尊重彼此的需要，彼此间相互学习和协作。在困难面前，在狩猎、采集和养育幼儿时，这种合作能力将被证明是至关重要的。因此这种家庭会留下更多的后代。作为个人，我们可以付出分担痛苦的代价，因为我们有同情苦难的能力，但我们获得了我们的文化和社会稳定所带来的全部好处。镜像神经元——及其馈赠的分享他人的情感的能力——不仅使我们能够与其他个体合拍，而且促使我们用它来深入了解善恶。我们欣慰地认识到，我们体内有某种东西，它使我们真正关心他人，就如同他们是我们自身扩展的一部分。

第3章
如何增强人自身

◎尼克·博斯特伦

尼克·博斯特伦（Nick Bostrom）

哲学家，牛津大学人类未来研究所所长。在到牛津大学之前，他曾在耶鲁大学教授哲学。他于2000年获伦敦经济学院博士学位。他具有物理学、计算神经科学和数理逻辑方面的背景，研究领域包括概率论基础、全球性灾害风险、人类增强伦理学（the ethics of human enhancement）和未来技术的影响等。

博斯特伦是《人类增强》一书的编者之一（另一位是朱利叶·萨乌莱斯），并与米兰·契尔科维奇共同主编了《全球性灾害风险》一书。他关于人类增强方面的论文有着广泛的影响力，他的著作已译成16种文字。他还曾短暂兼任过一些机构的顾问，包括欧洲委员会和美国中央情报局。

自然的智慧

医学是困难的。我们知道，这是因为尽管有时我们尽了最大努力，但往往还是以失败告终。然而，药物通常只是为了医治机体损伤，而体质增强则与此不同，它是针对未受损伤的机体系统，是要使它变得更强健——从多方面来看，这都是一个更宏大的目标。

人是进化复杂性的一个奇迹。虽然生物科学已有了很大进步，但这种复杂性的大部分仍未被揭示开来。当我们对这一复杂的、进化了的、知之甚少的系统进行医治时，我们的干预往往失败或适得其反。现已证明，止痛药会导致新生儿先天缺陷或高血压；号称安全的杀虫剂被发现具有致癌作用；发明的婴儿配方奶粉被吹嘘得优于母乳，但后来发现，

它缺少对大脑发育具有重要意义的某些脂肪酸。这种失败的事例可以罗列得很长很长。医源性疾病，即由医疗失误造成的恶化，据估计每年造成全美近22.5万人死亡，成为第三大死因。就好像存在一种“自然的智慧”，我们胆敢忽视它或凌驾其上，就会带来危险。对大自然智慧的笃信——或者反过来，对精心篡改自然尤其是篡改人性能够取得成功的怀疑——往往表现为对增强体质的做法持道义上的反对。很多人直觉上就认为“自然”是不可超越的，人类的狂妄自大终将付出代价。有些人可能会将这些想法归于神学教义的基础，但实际上并不存在这样的基础，有的只是对改变现状的不适。

人类体质增强的热心倡导者——对通过生物医学干预来改善认知、丰富情感或延缓衰老抱有积极乐观态度的人——总觉得所谓“自然的”偏好是纯粹的迷信，或无奈地得出结论认为，观点上的区别在于基本道德上的分歧，要对这种道德进行理性的讨论是不可能的。但是，我们做的可以比这更好。我们能够认识到，大自然有一定的“智慧”这种想法是有道理的。通过更好地了解这种智慧的程度和局限性，我们可以为增强人类体质的前景建立一套切实可行的启发性研究模式或经验法则。这种启发模式不是要替代通常的医学研究——譬如临床试验等——而是可以通过识别干预参数来引导研究，尽可能使干预带来的潜在的副作用和风险变得一目了然，这样一旦出现这类警告，我们就会格外谨慎行事。

通过考虑我们这个物种的起源和进化过程中的缺陷，我们可以更容易地看出大自然在哪个地方做得不够好。在这种地方，也只有在这种地方，采取可行的人类体质增强措施才比较容易。

进化——一个伟大的工程师

我们可将进化比喻为一位技巧娴熟的工程师。这一比喻的界限设在对我们有用的地方。

进化造就了人体组织这一系统。这是一个比我们人类迄今所构建的任何结构都要复杂得多的系统。我们惊叹人体组织的复杂性，它的各个器官在解决复杂问题方面进化得可谓完善：眼睛用来收集和预处理图像信息，免疫系统可以抵抗感染和癌症，肺部为血液提供氧气。人类的大脑——各种增强体质建议中最具吸引力的集中体现——可以说是已知宇宙中最复杂的对象。

鉴于我们对人体结构特别是大脑已有初步了解，我们希望如何来进一步强化这一系统呢？这相当于要超越进化。我们可能会怀疑，依据我们目前的工具和科学认识水平，我们是否有能力做到这一点。我建议我们不妨将这种模糊的疑虑变成一个问题，这也是提出任何强化干预措施时应当提出的问题。我们可以称之为最优进化挑战：如果拟议的干预能够导致体质增强，为什么我们没有发展成为那样？

可能的答案有三种：

·权衡利弊的标准发生了改变；

·价值取向不一致；

·进化限制。

如果上述任何一种答案能够为为什么进化没有形成所需的增强这一问题提供一种解释，那么增强性干预就可被认为是有希望的。当然它仍然需要临床试验，但这种启发模型为我们进入这个阶段开了绿灯。相反，如果不存在这样的答案，那么这种增强体质的建议很可能归于失败：它们可能根本行不通，也可能有严重的副作用（这种副作用当时可能不那么明显，需要有数年时间才能体现）。在这种情况下，启发模型警告我们须格外小心，或尝试其他方法。

权衡利弊的标准发生了改变

人体组织是在特定环境下进化的。这种特定环境就是非洲热带草原上的狩猎采集生活。现在，这个系统必须适应在现代世界的环境下工作，这是一个完全不同的环境。现代生存条件对我们这个物种来说来得太快，我们很难完全适应。因此进化所达成的取舍可能不再是最佳的。我们可以做出一些调整和改变来使人类机体更好地适应新的环境，尽管与创造最初设计的进化过程比起来，我们的工程改造过程要短得多。

对发达国家的大多数人来说，现代生活的一个新面貌是食物供应非常充分，人已摆脱了对季节的依赖。相比之下，在自然状态下，粮食在大部分时间里是相当稀缺的，人类不得不将能量储备放在第一位，并不得不采取高新陈代谢的机体组织、生命过程和行为等高耗能举措来维持

生存。例如，人类的大脑只占体重的2%，而其能耗却超过人体全部能耗的20%（对于新生儿，脑代谢甚至占到总代谢的60%）。脑、心、胃肠道、肾脏和肝脏消耗掉基础代谢的70%。进化不得不作出艰难的取舍，以便在这些器官的大小和功能方面，与在寻找食物和营养价值最大化之外，进行其他活动所花费的时间和精力方面之间取得平衡。

假设存在某种潜在的增进人类体质的途径，使我们拥有更多的智能。现在我们将其应用到这样一种问题上："如果有这样的好事，为什么我们没有自然地进化出这样一种高智能水平呢？"一种可能的答案是，进化改变了这种取舍权衡。因为智能水平越高（也许本身是有益的），所需的新陈代谢支出就越大（精神活动需要大脑消耗更多的能量）。当然，今天我们不再关心如何维持身体所需的卡路里（如果这是个问题的话，我们所关心的正好相反）。因此，如果我们可以找到某种方法，既能增进智能，又允许机体有更多的热量消耗，那么也许就可以达成一种新的绝佳平衡。这里，改变这种权衡的最佳点是靠增加粮食供应来保证的。实际上我们已经有了这样的方法：譬如咖啡因和莫达非尼等兴奋剂，它们对提高智能有效，同时副作用也相对较小。

也许大自然还可以设计出这样一种增进智力的遗传机制，其代价是使得头变得更大或成熟期相对延长。我们很容易看出为什么大自然没按此演化，即使它对增进智力有效。因为高的认知能力可能使狩猎采集者具有优势，但过大的头部或较长的成熟期带来的风险则会超过这种收益。但现在，随着新的生存手段的实现，譬如助产技术、剖腹产和儿童成长的安全环境，这些负担要小得多。

另一个潜在的增强我们机体的手段可能是提高细胞DNA的修复能力。这将有助于预防癌症和衰老。如果增加的成本仅仅是要求提高卡路里的摄取量，那这将是一个完全值得的代价。

除了提供新的资源，生存环境也会因向我们提出新的要求而改变。在书面语言发明之前，当然更谈不上文字的选择。识数、分析能力、抽象思维和长时间集中于某项认知挑战的能力在当今世界要比在更新世（冰川世）有用得多。这些需求上的变化也意味着可能的改进。例如，我们可能会发现某种药物能够提高我们的注意力。注意力既然这么有用，为什么我们不曾拥有一种更强健的集中注意力的能力？这或许也涉及取舍平衡问题。也许高强度集中注意力会增加大脑的新陈代谢，或减

少对周边的警觉。如今，卡路里不再稀缺，而对周边的警觉对猎人才显得非常重要（小心狮子就在你身后）。在现代社会中，更重要的是如何将注意力集中在（譬如）书本上，或计算机屏幕上，或与之交流的对话者上。意料之中的是，有些人是靠如尼古丁或利他林这类东西来保持注意力集中的。这些药物就是对最优进化挑战的回应。

价值取向不一致

对最优进化挑战的第二类答案源自于我们希望采用的标准与控制进化的标准之间的偏差。进化选择讲究内在适合度（inclusive fitness），就是说，这种适应性既包括个体自身的适应性，也包括由基因亲缘度所决定的社会伙伴的适应性。而我们在优化方面看重的则是这样一些价值：健康、成就、有知识、有意义的关系和美德——对我们来说，这些品质要比尽最大可能繁育后代更重要。

这种目标上的背离为有前途的体质增强提供了丰富的来源。在我们能指出这类价值取向不一致的地方，我们不必假定人类的才智超越自然进化就可以应对最优进化的挑战。一个从没有设计过汽车的普通技术员（更不用说那些精英了）就能够将一辆宝马改造成一个粗陋的雨水收集装置。如果碰巧我们这会儿需要的正是集雨器，那它的价值就超过一辆汽车，从我们的观点看，这也应算是一种增强。同样，我们或许可以改善人的机体，以便更好地服务于我们的特异目标，尽管我们无法像自然界的生存繁育机制那样来增强其表现。

这并非对我们希望增强的所有生物性状都有效。我们看重的许多性状在人类早期的生存环境中也起着提高繁殖成功的作用，健康就是一个例子。但我们看重的许多其他性状则对进化的成功没有贡献。避孕技术就是一个明显的例子。输精管结扎术、避孕药和其他避孕方法可看作是一类增强措施，因为它们能够提高我们对生殖系统的控制。我们看重这类增强措施，是因为它们促进了计划生育并提高了择偶的选择性。为什么进化没有为我们提供一种简单的生殖开关功能，这并不难理解。

我们至少可以区分两种不同形式的价值取向偏离。其一，那种使个体的内在适合度最大化的性状未必总是其最优性状。其二，这些性状对社会或非个体对象来说也未必总是最好的。从某种角度来看，这两种价值取向的不一致性表明增强型干预有可能是可行的。

例如，我们可以采取干预措施来提高我们的厌烦阈值水平。这种干预可能会使我们生活得更加愉快：我们可能会发现更多的令人感兴趣的事情，并从工作中、业余爱好中和其他人身上得到更多的快乐。按理说，低的厌烦阈值是人类对早期生存环境的一种适应，它会阻止我们将时间和精力浪费在与生存和繁衍生息无关的事情上。这里不是说我们应该将我们的耐烦能力提到无限高（即看什么都不烦，都高兴——译者注），而是说平均来看，我们的厌烦阈值应高于让我们感到生活最美好和满足的那些事情的水平。通过某种药物或基因干预的手段，适当提高这种厌烦阈值是可行的。

我们还可以考虑那些有益于社会或整个人类的潜在的增强手段。进化为什么没提供给我们这方面的优势（例如，扩展利他主义以及控制暴力冲动的能力），其原因是不言自明的。我们很容易看出进化是如何通过自然选择来使这些性状保持在一定水平上，从而使个体的内在适合度达到最大化。我们也看到，为什么当今社会可以从（平均而言）个人具有更大程度的利他主义和更好的控制暴力冲动的能力上受益。增强这些品质对现代人来说既是可行的也是社会的需要，其途径包括服用具有移情作用的药物（empathogenic drugs）、情绪疏导、提高修养，以及增强这些亲社会倾向的文化和社会环境。

即使是我们希望增强的某种品质本身具有促进内在适合度的效果，有时我们仍会借助于价值取向偏差来应对最优进化的挑战。譬如下述情景：一种性状（品质）可能内在地就关联着另一种具有降低适应性水平的性状（品质）。进化已经在这两种性状之间达成某种取舍平衡，但如果我们对这两种品质评价的优劣次序与进化的结果不同，我们就有理由采取不同的取舍。例如，假设非常高的创造性智力需要以注意力分散为代价。虽然我们在其他条件相同的情形下通常不愿意分散注意力，但如果注意力分散能带来很高的知识创新能力，那么我们可能会愿意接受相当程度的注意力分散。如果我们能够证明，在早期人类环境下这种偏好的取舍会使得适应性变差，那我们就有理由认为，今天我们应当寻求某种干预来得到我们想要的结果。与此相反，如果一种拟议的干预能给我们提供好处却不需要付出任何代价，我们就该怀疑这种干预——除非我们能在回应最优进化挑战的启发式模型的另两种答案中找到一个答案。

进化限制

回应最优进化挑战的最后一种潜在答案在理论上要比上述两种复杂一些。它源于这样一个事实：进化不是一个完美的适应性优化的过程。进化的优化要受到一整套重要的限制条件的约束。在一些特定情形下，这些限制条件意味着我们能够做得比自然进化过程更好。我们继续拿汽车作比喻：尽管我们的工程师比不上原汽车厂家的工程师水平，但只要我们有更好的工具和材料，我们就能够生产出更好的汽车。

进化限制条件可分为三类：

·毫无可能（进化根本不能产生特定的性状）；

·陷入困境（进化停留在不包含该性状的“局部最优”）；

·进化滞后（性状进化需要的时间非常之长，其更替的代数超出整个人类人口繁衍所用的时间）。

首先，我们来考虑“毫无可能”。生命过程只能出现在允许它存在的地方。例如，陆地生物不可能结晶出金刚石或大的金属物，因此进化不可能做到演变出钻石牙釉质或钛金属骨架。如果这些改进手段在技术上是可行的，那么它们很容易应对最优进化的挑战。这样的例子太多了。进化不可能生产出高性能硅芯片以加强神经网络计算，尽管这种芯片能提供明显的益处。虽然这种改进提出了许多技术上的挑战，譬如如何确保植入体内的芯片能克服机体排异性，但这种增强是不是行得通，这既不构成特定的进化原因，也与自然的智慧不相关。

所以会出现陷入局部最优困境的局面，是因为自然选择是一种“短视的”搜索过程，它会滞留在这样一种状态，尽管大的变动可能会使整个状态变得更好，但任何微小的变化都会使该状态变糟，于是这种状态就不会轻易改变。这方面的一个例子是人的阑尾。阑尾是我们食草的灵长类祖先具有的大得多的盲肠的残留物。尽管阑尾可能有某种有限的免疫功能，但它很容易受到感染。在自然状态下，阑尾炎是一种危及生命的疾病，尤其是在青少年身上发生的可能性很高。按说阑尾的退化很可能会增强早期人类对环境的适应性，然而，较小的阑尾会增加罹患阑尾炎的危险，于是那些小阑尾的基因携带者就会比非小阑尾携带者有更大的罹患阑尾炎的风险，因此，其适应性就较低。

除非进化能够找到一种一次性消除阑尾的方式，否则就无法摆脱这

个器官，因此它到现在仍然存在于我们体内。安全方便地切除阑尾作为一种干预或许是一种增强体质的手段，这样既提高了身体素质也改善了生活质量。

进化锁定在次优状态的另一种方式是杂合优势现象。这是指这样一种常见的情形：那些携带特定的异质结合基因的个体（即携带两个不同版本基因的个体）较携带纯合基因的个体（拥有两套完全相同副本的个体）有优势。在许多情况下，潜在的有害基因之所以能够在一定人群中维持有限的频率，杂合优势是一个重要原因。一个典型的例子是不完全隐性镰状细胞基因：纯合子个体易患镰状细胞贫血症（一种潜在致命的血液性疾病），而杂合子个体在这种情形下就显示出优势：提高了抵御患上疟疾的免疫力。镰状细胞杂合子携带者要比两种类型的纯合子（一种是缺乏镰状细胞基因，另一种是有两套副本）携带者有更大的适应性。平衡选择以一定频率在人群中保留了镰状细胞基因，其分布恰好与疟疾流行的地理特征相一致。所谓“优化的”进化选择是指，某些人恰巧生来就带有纯合子等位基因，因而导致镰状细胞贫血症。所谓“理想优化”——每个人都携带杂合等位基因——是自然选择无法实现的，因为按照孟德尔遗传模型，杂合子父母生育的新生儿肯定有25%的可能是镰状细胞基因纯合子的携带者。

杂合子优势意味着一种明显的增强体质的机会。如果可能，变异的等位基因可以被删除掉，而且其基因产物可用作药物。另外，通过对体外受精的基因进行筛查可以保证杂合性，使我们能够取得自然选择未能实现的理想优化。

最后我们来讨论“进化滞后”——进化的优化能力的最终限制。进化的适应需要时间——很长一段时间。如环境条件变化得过快，基因的演化就会跟不上。人类祖先的生存条件是频繁变化的，譬如迁移到新的地区、气候变化、社会动荡、采用了先进工具，并且适应了与病原体、寄生虫、捕食者和猎物共存等等，人类从来没有完全适应其生存环境。进化不断地驱动着人类在适应性斜坡上奔跑，但适应性的景观在不断更新，人类的进化可能永远不会到达峰顶。即使存在有利的基因或基因组合，它们也没有时间来扩散到所有人群。

如果我们找到一种理想的、进化还没来得及将其扩散到整个人类的基因，那么通过干预将其插入基因组，或模仿其效应，就可能实现应对

最优进化的挑战。一个简单的例子是乳糖耐受性。乳糖不耐性曾是哺乳动物对促进断奶的一种适应性，但在过去5000年到1万年的时间里，乳制品一直刺激着人体对乳糖酶的选择，从而使人体发展出乳糖耐受性。但这段时间还没长到足以让这一性状扩散到所有人群，因此医疗上需要通过服用乳糖酶药丸来帮助患有乳糖不耐症的人群消化乳糖，这同时也扩大了这一人群的食谱。这种增强显然很好地回应了最优进化的挑战。也许我们还可以找到类似的情况：譬如参与脑发育的基因也受到强烈的正向选择，在过去的3.7万年（甚至可能更近）里已经出现了新的变种，这个时间段从进化上看还是相当短的。

结论：进化的缺陷应该看作是大自然赋予我们的机会。系统研究曾创造人类机体的进化过程的局限性，我们就能够采用干预措施来找出增强人类机体的有前途的方法，这些干预措施有些在今天已是可行的，有些则在不远的将来有可能实现。

从长远来看，这只是第一步，我们应该能够走得更远。一旦我们对人的机体有了比较完整的了解，或者换句话说，我们学会了如何构造出具有同样复杂程度和性能的完全人工系统，那么我们将不再需要借助于进化启发模型和“自然的智慧”的支持。总有一天，我们会学会如何从头设计出新的器官和机体。我们甚至可以不再依赖于生物材料来行使我们的身体和大脑的功能。人类一旦从这些实际的限制中解放出来，其任务将变成如何明智地运用我们的力量来自我完善。换句话说，我们面临的挑战将从目前的科学实践为主转变为道德实践为主。如果说，从我们目前的观点看，道德任务似乎还显得相对琐碎的话，这只能说它反映了我们目前的不成熟。

第4章
我们在非自然宇宙中的位置

◎肖恩·卡罗尔

肖恩·卡罗尔（Sean Carroll）

自2006年以来，一直担任加州理工学院高级研究助理，1993年在哈佛大学获物理学博士学位，论文题目是“场论中拓扑和几何现象的宇宙学结果”。他曾先后在麻省理工学院理论物理中心和加州圣芭芭拉大学理论物理研究所做博士后研究，任芝加哥大学物理学助教。他的研究范围横跨理论物理学多个方向，包括宇宙学、场论、粒子物理学和引力理论。他目前的研究方向是暗物质和暗能量的性质；宇宙学、量子引力和统计力学之间的联系以及早期宇宙是否经历了一个暴胀期。

卡罗尔撰写过研究生教材《时空和几何：广义相对论引论》，并为教学出版公司录制过有关暗物质和暗能量方面的课程。他曾分别获得过斯隆基金会（Alfred P.Sloan Foundation）、戴维和露西尔帕卡德基金会（David and Lucile Packard Foundation）提供的奖学金。获得的荣誉还包括有麻省理工学院研究生会教学奖和维拉诺瓦大学艺术与科学学院奖章，2007年被授予美国国家科学基金会NSF杰出讲师称号。他是宇宙变化博客http：//blogs.discovermagazine.com/cosmicvariance/的创始人之一和博主。

宇宙一直想告诉我们一些东西。但到目前为止，我们始终没搞清楚它要说的是什么。

科学是要增进理解，但科学的进步却往往由误解推进。当一个接一个的观察与我们的预言相一致时，科学就很难再向前推进了。但当实验结果与我们钟爱的理论背道而驰时，我们便开始取得某些进展。再好的

理论也不可能涵盖我们希望得到解释的一切现象。我们的目标是要将理论扩大到未知范围，没有什么能比与当前理论图像相矛盾的事实更有用的了。

但是，我们并不总是这么幸运。有些理论，在其适用范围内，对我们的观测结果能够给予很好的解释，但让我们有一种令人不安的感觉，就是它们不是事情的最终解决方案。这方面的一个绝好的例子是粒子物理学的标准模型。经过20世纪60～70年代的辛勤建构，这个模型在解释粒子物理学家在20世纪80～90年代积累起来的海量数据方面发挥了英雄般的作用。直到最近——随着中微子质量和存在暗物质和暗能量证据的发现——标准模型才开始遭遇挫折，我们至今仍不知道可以用什么新的理论来取代它。

可是在粒子物理学领域，从来就没有人真的认为标准模型是最终解决方案。一方面，它不包括基本力之一——引力——因此它被认为不可能是完备的。另一方面，它似乎不像是完全“自然的”。它包含了大量的粒子，要用很多参数来描述，而且这些参数值似乎显得杂乱无章，毫无规律。它有各种不同的对称性——有些破缺，有些不破缺——但都没有令人信服的理由。对于弱核力，能量上还存在一个很重要的尺度，如果你不做测量，你想象不到这个尺度会有多小。

尽管标准模型勉强经受住了实验的检验，但它总是不太对味儿。在这一杂乱无章的理论背后一定有一些更为简单、更为牢靠的安排。

旁观者清。在有些人看起来非常自然的东西在另一些人看来则大有疑问。但是物理学家在判定一个理论是否自然时有着非常明确的标准。任何好的物理理论都有一组数——自然常数，至少在特定模型的适用范围内是如此。如果一个或多个这样的常数过大或过小，我们会认为这个模型不够自然。尤其是我们不知道大自然是否给了我们一些线索，让我们能够了解所看到现象背后的物理，即我们据以解释这些过大或过小常数的更简单的概念。

不自然常数的一个明显的例子是真空能——真空空间的能量密度。我们习惯于认为能量与某种物质形式——粒子或辐射或运动——相联系。但在广义相对论——爱因斯坦的时空理论——里，能量是空间本身所固有的结构。我们怎么知道这一点的呢？因为根据爱因斯坦理论，每

一种能量形式都对空间的扩展有贡献。真空能——其密度保持不变，而普通物质和辐射则是耗散性的——对宇宙的膨胀给予持续的冲击，导致遥远的星系离我们加速而去。实际上在1998年我们就已检测到宇宙膨胀的这种加速现象，它表明真空能的这一假说性概念是真实存在的。

但它不是那么自然。问题不在于真空能概念本身，而是真空能的大小不足以解释这个加速度：它远远小于应具有的值。我们可以将自然界其他常数——光速、牛顿的万有引力常数、量子力学的普朗克常数——的测量值组合起来来预测真空能应具有的值。结果发现这个值约为10^{112} erg/cm^{-3} 。如果说它看起来像个大数，那算没说错，因为测量值只有约10^{-8} erg/cm^{-3} 。两个原本应该相等的数相差了10^{120} 倍（10后面跟有120个0）。这是怎么回事？

也许宇宙一直在告诉我们一些东西。观察到的真空能似乎小得很不自然，但顺便说一下：如果这个值大很多，那我们就不可能在这里观察它了。如果真空能的值接近于其“自然”值10^{112} erg/cm^{-3} ，那么空间加速膨胀的惊人速度将使单个原子都会被撕裂，就更不用说行星、恒星和星系了。

因此也许我们只是运气好，或者说也许是运气相当不好——我们观察到的真空能的值之所以非常之小，就因为我们生活在宇宙的一个不寻常的部分里。

可观测宇宙——我们看到的我们周围的那部分——在非常大的尺度上似乎显得极为相似。在数十亿光年的空间区域内，星系密度差不多相同。但是我们看不到整个宇宙，因为光传播的速度有限，因此我们可观察的范围有一个限度。在可观察视野之外，宇宙可能会无限伸展，未必是统一的。事实上，局部得出的物理学定律和自然常数可能会逐渐不同。想象一个打满补丁的宇宙——在每块补丁之内，条件和参数是统一的；但各块补丁本身则彼此完全不同。我们会发现，我们只是处在一块好客的补丁上。也许真空能的值在我们看来显得不正常是因为我们看到的只是整体的一个明显不具代表性的部分，宇宙从整体上说就像任何人希望的那样是非常自然的。

或者……也许不是。这种人存推理——试图根据由更大的、不可观测的“多元宇宙”理论得出的选择效应来解释可观测宇宙的不寻常特征

——使得很多人感到不舒服，如果不说是完全反对的话。一方面，大量引用不同层次物理学局部定律来解释区区几个数字似乎并不十分经济。其次，我们很难知道这些推理是否正确。多元宇宙看似解释了为什么这样的条件允许我们存在，但问题是是否存在这样的多元宇宙。如果自然常数不符合我们的存在，我们同样不可能在这里深入思考这些事情。具有不同局部物理学定律的补丁空间拼成的宇宙可以给自然参数的不自然一个解释；还有一种解释是，物理学定律是独一无二的——而且恰好与智能生命的存在兼容。在这两种解释里，我们到底该选择哪一个？

我们这个可观测宇宙的另一个不寻常特征——大爆炸——提供了一种可能的线索。宇宙学家们在两种不同意义上运用这句话。大爆炸本身是指宇宙诞生的奇异时刻，而大爆炸模型则是一种假设，用来描述我们的宇宙从最初的炽热、致密状态到我们今天看到的冷却了的、膨胀着的状态的演化。对宇宙史上极早期所发生的事情我们知道得很少，因此对大爆炸时刻的描述纯属推测，但宇宙从极早期到今天的演化是相当清楚的。因此大爆炸模型有坚实的经验基础，尽管大爆炸本身对我们来说还只是个占位符。然而，正如粒子物理学里的标准模型那样，大爆炸模型在数据拟合上的成功并不意味着它在自然性方面不存在棘手的问题。

现代宇宙学里不为公众所知的难题之一是，我们不知道为什么早期宇宙看上去是那个样子。这里我们谈论的不是真空能问题里的自然常数，而是有关宇宙自身的结构。目前，这种结构看起来就像是一个由气体、恒星和暗物质（一种尚未被直接观测到，但其引力对银河系结构具有影响的物质形式）构成的稀薄、弥散的集合体，这种集合体聚合成散布在整个宇宙中的星系，并在真空能背景下演化。所有东西都处在不断膨胀和降温过程中，因此宇宙在过去要比现在更热、更致密。宇宙在过去也较为平滑，它是从几乎完全均匀的状态下开始演化的。自从140亿年前的大爆炸以来，物质受到引力无情的拖拽，逐渐聚集成恒星和星系。就我们目前所知，宇宙还将继续膨胀下去。星系将分离得越来越远，当所有物质弥散到虚空之后，宇宙最终将再一次变得完全平滑。在遥远的未来，宇宙将是一锅基本粒子稀粥，所有粒子将变得越来越冷，分离得越来越远。

这个故事的一个十分明显的特点是：宇宙的过去非常不同于未来。早期宇宙是炽热和致密的，而晚期宇宙则是寒冷和稀薄的。好……那为什么会这样呢？事实是，我们不知道。

我们对宇宙的过去和未来之间区分的认识是如此根深蒂固，以至于似乎并不需要多做解释。像鱼儿无视水的存在，我们几乎从不在意时间不对称——所谓时间之箭——这一深刻谜团。但它却是我们局部物理环境的最突出的特点。我们可以将鸡蛋变成煎蛋卷，但我们不能把煎蛋卷变回鸡蛋。我们可以记住过去但不能记住未来。任何看过电影倒片的人都知道，时间根本不可能倒转，否则事情很快就会变得荒诞不经。在我们的局部环境动力学里，过去和将来之间不存在对称性。令人费解的是，物理学基本定律却存在这样的对称性；它们在时间上向前和向后同样有效。如果物理定律告诉我们过去和将来都是平等的，为什么我们的日常经验却告诉我们事情并非都如此？

时间箭头的起源，可以从我们早上吃早餐时打一个鸡蛋一直追溯到宇宙的开端。物理学家是根据熵——对物理状态混乱度的一种量度——来跟踪时间流逝的。如果你将某个物体集合分散开来，它们可以有任意多种无序的安排——对此我们指定一个高熵值。但要将这些物体精确地排列起来则只有少数几种方法，这种有序排列的熵相对较低。一副有序的纸牌具有低熵，而洗好的牌则具有很高的熵。高熵结构，如果你愿意的话，要比低熵结构更自然，原因很简单，存在的高熵结构要多得多。

由于高熵结构更自然，数量也更多，因此孤立物理系统的熵总是随着时间推移趋于增加（或至少不减少）。这就是热力学第二定律，它也许是物理学最值得珍视的原理。第二定律确保了煎蛋卷不会自发地变成鸡蛋，鸡蛋的熵远低于煎蛋卷，因为组成鸡蛋的分子排列方式要比煎蛋卷的分子排列方式少得多。

但是，这只是故事的一半。虽然熵趋向于增加，因为高熵比低熵有更多的存在方式，但这并不能解释为什么当初熵很低。但事实就是这样。我们的宇宙的熵开始时非常小，此后不断在增大。想想看：我们的可观测宇宙在超过10亿光年的范围内包含了大量的粒子（准确地说，约为10^{88} 个）。然而在早期，它们都以高温稠密等离子体形式小心地挤在非常狭小的空间区域里。这该多不自然！所有这些粒子彼此相距甚远地分散开来的方式要多得多，而这正是随着时间推移所发生的事情。与宇宙熵可能的取值相比，我们的宇宙目前的熵值处于中间。未来，熵将变得巨大，因为它显然更愿意这样，但在过去，熵值却非常小。

不仅是某些自然常数明显受到过微调，而且宇宙的早期状态也如

此。宇宙为什么不一直处在一个高熵状态？没有明显的理由说明宇宙中的物质和辐射曾经非要被挤压在一起，它同样可以是永远处于稀薄弥散状态。但那样就没有时间箭头了，宇宙只能坐在那里，过去和未来没有任何区别，也不可能发生任何事情。

你可以想象一下这将会怎样。在一个真正高熵的宇宙里，什么都不发生，这其中就包含我们的生命。我们用来刻画生命的每一个特征——代谢、繁殖、进化、记忆——很大程度上都取决于时间箭头。生命，可以毫不夸张地说，是由熵增在时间上推进的。如果熵一直很大，也不增长，就不会有生命。我们可以用人存推理来解释大爆炸时的低熵吗？

不尽然。尽管存在人存推理本质上无法验证的担心，但它确实具有一定的逻辑：如果某种不自然的事情是生命存在所必需的，那么我们就应当观察到它经过足够的微调以便说明好客的宇宙，此外别无其他。这正是我们从真空能的观测值里发现的东西：如果它哪怕大一点点，我们就不会在这里谈论它了。

但熵的情况不同。早期宇宙不仅是一种非常特殊的结构，而且它的熵远低于用来说明我们的存在所需的值。我们的存在充其量也只需要我们在银河系里发现的环境。但银河系只是我们可观测宇宙中千亿个星系中的一个，那些所有的星系又以什么理由存在呢？特别是，为什么聚集成这些星系的物质不像真正的高熵宇宙那样均匀散布在整个空间呢？即使生命的存在需要宇宙有一个低熵的开端，但我们的实际宇宙在早期的微调值似乎没必要如此挥霍。

现在我们已经进入这样一种境地：我们不仅缺乏得到实验结果广泛支持的理论，甚至没有精心制定的假说。尽管如此，越来越多的宇宙学家正在认真对待大爆炸之前发生的问题。请记住，我们对大爆炸本身并不充分了解；我们倾向于把它看成是具有无限大密度和曲率的时刻，但事实是我们根本就不知道该怎么做。我们不知道该如何协调量子力学要求与广义相对论弯曲时空之间的关系。一旦我们知道该如何协调了，那么所谓奇点和禁闭等问题就可能通过变得平稳和连续而得到解决。越来越多的物理学家敢于对大爆炸之前的事情发挥想象力，认为我们的可观测宇宙可能有一个不可观察的史前史。

在这一史前时期，我们可能会找到对早期宇宙的微调过的低熵状态

的一个解释。想象一个处于据称高熵状态的宇宙：寒冷而稀薄，所有粒子彼此分得很开。但要记住，还有真空能，因此，即使是虚空空间，也不是完全平静的。量子力学法则告诉我们，只要存在这种能量，真空就存在涨落。粒子不时地从无到有又从有到无，各种场时而显出统计上很难出现的结构。如果我们等待足够长时间，那么恰当的结构就会出现——这就是产生一种全新宇宙的必要条件。物质和能量的短暂涨落可能会积聚于某个小区域，使得时空结构扭曲得恰好让空间与时间断开，并产生一个不连续的空间泡泡。这个泡泡能够膨胀和发展，最终冷却下来并冷凝成恒星和星系。这可能就是我们生活其中的宇宙。

换句话说，借助于大爆炸前的条件，我们或许能够解释大爆炸时的明显微调了的特征。也许我们看到的所有粒子最初都处在一个狭小致密的区域里，因为这样的结构更容易产生一个新的泡泡宇宙，而不是一开始就产生一个大的、稀薄的宇宙。我们的可观测宇宙的熵增以及相应的时间箭头可能是更大的多元宇宙无法通过产生新的婴儿宇宙来满足创造更大熵的欲望的一种反映。如果我们能够一览整个宇宙结构，那么一切看起来可能都很自然。

或者……也许不是这样。我们还是缺乏足够的资金将这种思想付诸检验。但有时我们必须孤立前行，提出并改进想法，然后才能够充分理解它们，知道如何对它们进行检验。在我们努力搞清楚这些想法会将我们引向何方之前，我们什么都不知道。宇宙当然想告诉我们一些东西，我们要做的就是竭尽所能搞清楚它说的东西。

第5章
暗能量到底是什么？

◎斯蒂芬·亚历山大

斯蒂芬·亚历山大（Stephen H.S.Alexander）

哈弗福德文理学院物理学副教授。2000年从布朗大学获得物理学博士学位，论文题目是“弦论和宇宙学之间的接口”。2000～2002年，他获得PPARC（英国粒子物理和天文学研究委员会）博士后资助，最近又获得了国家科学基金事业奖，并当选为国家地理新兴探索者。他的研究集中于尚未解决的问题——譬如宇宙常数或暗能量等问题——这些问题将宇宙论、量子引力和基本粒子的标准模型联系在一起。特别是，他将宇宙学观测运用到构建和检验基本理论上。

物理学家知道，一切自然现象——以及大部分先进技术，如手机和全球定位卫星——都可以用两种物理学原理——量子力学和相对论——来给予解释。最初我们看不出这两种原理之间的联系，但现在通过强有力的和令人惊讶的方式已经证明，事实确实如此。在20世纪20年代后期，保罗·狄拉克和其他一些人发展了量子场的概念，它显示了如何将爱因斯坦的狭义相对论与量子力学结合起来来解释自然界中4种基本力中的3种——电磁力、弱作用力和强作用力。

尽管有这样的成功，但量子力学与爱因斯坦广义相对论（描述第4种基本力——引力）的统一则尚未解决。学物理的学生一旦对广义相对论和量子力学的深远影响和魔力变得信心满满，他们就不禁会想象两者之间存在深刻联系。一些雄心勃勃的学生甚至可能决定将试图统一所有4种力作为自己的职业生涯。但是在本文里，我将说明——不谈美学——物理学界之所以被迫寻求统一是因为存在一种称为暗能量的看不见的奇异能量形式，它对广义相对论和量子力学均有深远影响。

今天的物理学状况非常像它在20世纪之交量子论和相对论曙光微露之前的情形。当时许多著名的物理学家认为，实验的目的只是彻底扫清细节，他们确信，掌管宇宙的物理学定律总的来说已经明了。美国第一个诺贝尔物理学奖得主阿尔伯特·迈克耳孙在1894年就明显持这种态度，他说："看来很可能宏大的（物理科学）基本原理已经牢固确立。"巧的是，还就得发明出量子力学和广义相对论才能解释这些"不起眼"的实验细节。面对暗能量问题的挑战，当今的物理学家应吸取历史教训，避免类似的假设。

迄今最为精确的物理理论当属量子电动力学（QED）。理查德·费恩曼的量子场论QED形式体系是量子力学和爱因斯坦狭义相对论的完美结合。量子电动力学方程可以给出由电子的量子自旋产生的细微的磁场值，它与实验测得的值之间的误差小到小数点后面9个零以后（相当于纽约和洛杉矶之间的距离测量精确到一根头发丝的粗细）。但其母理论——量子场论（QFT），它将空间尺度小到万亿分之一厘米，大到复杂分子行为的（除引力之外）所有的物理学统一起来——则要求存在这样一种能量形式——暗能量，并且所做的预言与观察的偏差将小于小数点后面120个0之后！

暗能量本身是无法直接观察到的，它是已知的最令人困惑的物质形式——是唯一一种其作用范围小到亚原子尺度，大到宇宙中最大距离的"东西"。我们在原子核深处和遥远的恒星运动中都发现了它活动的证据。它的普适性超越了包括爱因斯坦在内的伟大头脑的想象力。因为它如此难以捉摸，故暗能量还有各式各样的别称，其中包括宇宙学常数和真空能（我会交替使用这些名称）。本文将展示存在于我们的宇宙中的最小尺度和最大尺度之间的惊人联系。

我希望你能将充满宇宙的虚空空间想象成海洋。我们知道，海面上的小波浪会发展成海啸。暗能量就是一种在空间结构上具有波浪特性的物质形式。在亚原子尺度上，这些暗能量波表现为量子现象，这一点我稍后再作讨论。我们先来描述在广袤的空间上，或者说在宇宙尺度上，在爱因斯坦的广义相对论里，暗能量是如何展现的。

相对论宇宙学

广义相对论告诉我们，空间和时间并不共同构成一个固定的、空洞

的物质和能量活动的舞台，相反，时空本身在对物质的反应中会变得弯曲。

空间为什么会有这样的表现？

爱因斯坦回答了这个问题。他向我们证明了用两种不同的方式——或者是加速，或者是引力吸引——都可以产生质量。让我们回顾一下爱因斯坦的思想实验：在两部分开的电梯里各有一个人，他们无法看到外面。一部电梯里的乘客并不知道自己是在外层空间。起初，他自由飘浮着，但一旦电梯开始加速上升，他的脚就会踏在电梯的地板上，让他感觉到重量（质量）。另一个人则待在静止在地球上的电梯里，具有相同的质量感觉，这是由于地球引力的缘故。这个想象的实验导致爱因斯坦提出了等效原理，他称之为“我一生中最快乐的想法”。等效原理是说，非引力环境下的加速运动等效于静止在一个引力场中。两种完全不同的运动状态给出同样的质量感觉。让我们探索由此带来的后果。

爱因斯坦意识到，在运动的相对性背后有一个新的实在，它向我们展示了引力性质的神秘性：质量和引力都是空间弯曲的表现。正如他的狭义相对论显示了在所有参考系里光速不变一样，广义相对论说明了所有的运动状态（包括静止状态）都是相对的，只有时空本身在存在物质或能量的情形下可以弯曲这一点是绝对的。

两部电梯里的乘客的经验可以用等效原理来解释。与运动电梯有关的能量使空间弯曲，从而使电梯加速。同样，地球的质量使空间弯曲，从而产生吸引性的引力。广义相对论的最终结果是，重力作用不需要背景空间，而是由它定义了背景空间。与此相对照，我们来考虑磁力，这种力是通过两个磁铁产生的磁场施加的，需要空间来扩展。同样，引力场具有引力，但性质上不同于其他的场，它不需要背景时空，它本身就是背景时空。

广义相对论适用于描述整个宇宙的空间和时间的行为，正像光速对于任何地方的观察者都相同一样。物理学家们试图用广义相对论来了解存在物质和能量情形下的时空结构和行为。已有的时空结构会影响到物质和光的运动。在某些情况下，譬如黑洞，空间可以弯曲到使得光无法逃脱。广义相对论在解释和预言引力的新特性方面取得了前所未有的成功。但仍有一个暗问题有待设法解决。

宇宙学思想实验

这里有一个思想实验。让我们将我们的宇宙空间看成是一个巨大的气球表面。其表面上的每个点均与所有其他各点相同。将每个星系想象成气球宇宙表面上对应的静态点。如果我们知道气球上的点的物质分布，那么广义相对论方程就会告诉我们这个表面空间是如何变化的。通过运用哥白尼原理的简单假设——宇宙中的物质是均匀分布的，不存在优先的观察者——和广义相对论，爱因斯坦和苏联数学家亚历山大·弗里德曼各自独立地找到了一种解释宇宙时空大尺度行为的解。他们发现，宇宙中的物质有一种吹气球的效应：使它膨胀。他们的方程解与哥白尼原理是相容的，但令人吃惊的是空间不具有静态性质（和观测偏差）。当时，宇宙的膨胀还没有被观测到，因此，他们的解似乎是一种错误的预言。

每个优秀的物理学家都知道在这种情况下如何拯救广义相对论——那就是插入"任意因子"以保留理论中好的部分，同时剔除不想要的部分。爱因斯坦认识到，他可以在他的方程中引入一个常数——宇宙学常数——来消除扩张或收缩，从而使宇宙保持稳恒态。1924年，天文学家埃德温·哈勃发现，宇宙确实在不断膨胀，于是爱因斯坦愉快地取消了这个宇宙学常数，但后来他将这一举动称为他一生中"最大的错误"。他哪里知道，这之后人们从物理上认识清楚了，这个任意因子的行为恰像一种不可见的、均匀充满整个宇宙的奇怪流体：这就是难以捉摸的暗能量。

爱因斯坦方程所描述的膨胀的宇宙准确地预言了宇宙的年龄为140亿岁。在零秒时刻，宇宙从一个奇点显现为一个致密的、辐射占主导地位的时代，随着宇宙的冷却，辐射蜕变成轻元素（从氢到锂），并由它们聚合成最初的恒星和星系。这种相对论宇宙学模型被称为标准的大爆炸图景，但是尽管它取得了巨大成功，但却不能解释哥白尼的均匀性原理为什么能够成为各种复杂结构——我们在夜空中看到的星系和星系群——的基础。但至少宇宙学常数问题暂时解决了。

后来，到了1998年，宇宙学家索尔·珀尔马特（Saul Perlmutter）、亚当·利斯（Adam Riess）和布赖恩·施密特（Brian Schmidt）望着遥远的所谓超新星爆发的恒星运动，意识到从哈勃观察得出的宇宙均匀膨胀的结论在大尺度上——这个尺度要比哈勃当时能够观测到的更大——是

不正确的。他们发现，宇宙膨胀呈加速状态。在试图解释什么样的东西会导致这种加速的过程中，宇宙学家很快找到了罪魁祸首：这就是宇宙学常数。事实上，宇宙学常数似乎是唯一既能够解释珀尔马特等人的观测结果，又能够说明诸如星系形成和维持等其他观测上难题的原因。宇宙学常数，现在更恰当的称呼叫暗能量，已不再是任意因子，它的真正来源和性质是当今整个物理学所面临的必须解决的问题。

在大尺度距离上，与引力比起来，电磁力和两种核力已无关紧要。现在我们有观测证据表明，暗能量——一种无所不在的、均匀分布的、排斥性的物质——作用于引力形成了空间加速的结构。问题是没有任何已知性质的其他物质具有这种使空间本身相斥的性质。暗能量具有所有亚原子粒子普遍具有的量子特性——天生就具有负压和正能量特性（负压可以看成是蹦床将孩子反弹到空中的推力）。

然而，广义相对论并不能令人满意地揭示出暗能量的这种性质及其来源。到目前为止，它能说的只是虚空空间可以有恒定的能量，这种恒定的能量能构成加速宇宙膨胀的斥力。尽管如此，广义相对论对有或没有宇宙学常数都可以满足了。我们对物质和能量的最准确的理解是在量子场论的亚原子领域。量子场论中的虚无或真空概念可以帮助确定这种恒定能量的来源。

感知场

实际上，比起其他东西来，我们更熟悉场。我们听到、看到、感受到它们。光、声和热用场来描述是最恰当的，因为它们在特定区域上具有一致性。由于其扩展性，自然界的所有4种基本力都是以场的形式存在的。例如，带电粒子向外辐射电场来吸引带相反电荷的粒子。爱因斯坦正是因为证明了光辐射可以从金属中打出电子而获得了诺贝尔奖，由此他建立了电磁场的类粒子性质。因此，场可以是粒子。但粒子可以是场吗？

另一种思想实验。想象一条河床在瀑布处中断的河。大部分下泄的水构成连续的瀑布，但总有些水滴溅出瀑布外。一小部分会蒸发掉，但大部分的水会在瀑布底端重新汇成溪流。量子场论就可以类比为瀑布。让我们考虑将电子本身作为一个场，这个场弥漫到空间任何地方。像水滴一样，电子最终会作为一个粒子回到它的母电磁场。

在量子力学与狭义相对论得到统一之后，量子场论做出了两个重要的预言。首先，场是存在于真空（即没有粒子的空间）中的基本客体。我们可以将河流看成是场的真空态。粒子可以从真空中产生或湮没，就像水滴可以离开随后又返回到瀑布。保罗·狄拉克发现量子场论有一个惊人的结果：反物质。他的理由是：既然真空状态是最低的可能的能量状态，因此粒子不可能自发产生，除非同时产生出其反粒子。

业已证明，真空中物质和反物质的自发冒泡会产生静电力，就是引起两片金属相互吸引的那种力。真空的这种量子特性称为卡西米尔效应，它是真空能存在的直接证据。现在我们用一句精辟的话概括一下：真空能量的性质其实就是暗能量的量子表现。当我们将这个能量看成是存在于虚空空间的每一个点上时，我们必然得出这样的结论：它对宇宙的贡献是一个接近于无限大的宇宙学常数，这将导致宇宙瞬间暴胀开来，其速度远远超出珀尔马特及其同事观察到的结果。根据量子场论，真空能是无限的，但我们可以在不影响核物理有用的预言前提下减去它，在核物理里量子场论可谓几乎完全有效。

在量子场论中减去几乎无限的真空能的可行性类似于降低海平面。如果你坐在海洋中的一条船上，你会感觉到海浪的上升和下降；但如果世界上的海洋在不断流失，这你就感觉不到了。你只能感觉得到作用在你船上的表面波。巨大的真空能就像大海的相对高度，可以被扣除掉而不会影响到量子场论的好的预言和在实验上的成功。我们在实验室测得的真空能的任何效应，就如同摇动船的波，需要将这种无限大的部分减去。

回想一下，广义相对论指出，所有能量和物质，无论其形式如何，都会对时空结构产生影响。因此暗能量也必然会作用在引力上，以极其巨大的速度产生越来越多的虚空空间。我们的宇宙目前的加速性质与这种预期是相符的。虽然量子场论——在亚原子物理学描述方面最为成功和技术上最有用的理论——告诉了我们为什么必然存在暗能量，但它对引力的预言则令人尴尬。我们还能说量子场论是完全正确的吗？我们能认为真空能的确几乎是无限的，只是引力对它不敏感吗？如果是这样的话，那么广义相对论里一定缺失了什么东西。另一种可能性是量子场论和广义相对论在描述暗能量和真空结构方面都是不正确的。

在过去的半个世纪里，物理学家们一直为暗能量/宇宙学常数问题

艰苦跋涉着。进展很是缓慢，虽然存在各种前景美妙的建议。我认为，我们必须公正地解决好广义相对论和量子化的两个基本原则才能找到解决方法。

新方向仍未可知

大多数被这个问题搞得灰头土脸的物理学家一致认为，要解决这个问题很可能需要我们对目前的图像进行重大改变。值得一提的是如下一些逻辑方向。

1.量子场论没错，需要改变的是引力

假定我们在亚原子尺度上测得了暗能量，那么我们可以得出结论说广义相对论是不足的，因为它的结构中没有将量子物理学包括进来。从量子力学上说，真空是一切物质和反物质的湮没。而在广义相对论看来，物质与反物质之间没有任何区别，实际上反物质的存在甚至没有意义，只有物质、能量和曲率三者之间的关系是有意义的——至少宏观上如此。但加拿大物理学家威廉·乔治·翁鲁（William George Unruh）证明，量子场论下的真空状态依赖于观察者的运动状态；在一个观察者看来是无物质的真空状态在另一个观察者看来却是一锅物质汤。因此量子场论真空与广义相对论是有冲突的，后者给予所有观察者的运动状态以平等的物理基础。

问题有可能是量子版的引力对量子场恒常活动产生的大量的暗能量自然地“视而不见”。理论物理学家一直试图对量子理论进行修改以使其与广义相对论相容，后者则对真空能“视而不见”。这种修改是要产生一种“屏蔽”效应，但所有这些更改在宇宙学距离上都出现了问题，我们在大尺度距离上观察到了暗能量的作用。

2.暗能量——表现得就像你的生命取决于它

通过用能量的振动弦取代基本的点状物质结构，弦论可以明显地将量子场论与引力统一起来。在弦论里，弦的数学解可以给出物质和能量的不同形式有关的不同的振动模式，这其中包括引力子——一种与引力相伴的粒子。但是，这种让人折服的终极统一图像是有代价的：那就是我们应该考虑的暗能量，我们自己在宇宙中的位置，甚至我们所说的宇

宙到底是什么意思，所有这些基本图像都可能需要改变。

除了将广义相对论和量子力学成功地统一起来之外，弦论还认为真空解可以有可数个无穷大，每一种给出多元宇宙中不同宇宙各自不同的暗能量值。一些权威弦论专家认为，如果这样的多元宇宙足够大，那么所有这些解都可以在空间遥远的某个部分实现，而我们就生活在其中一个部分里。诺贝尔奖获得者史蒂芬·温伯格证明，如果暗能量过大，它的斥性就会在大爆炸之后瞬间阻止轻元素坍缩成恒星和星系。然而事实是它们做到了这一点。这种论证就好比人存原理：我们恰好生活在多元宇宙的这样一个部分里，它的暗能量值恰好允许生命存在。这一原理可能为弦论找到一个归宿，并将改变我们对真空的认识——不同的真空可能会产生具有不同物理定律的不同宇宙。弦论家和数学家正在为实现这项建议而努力工作。

未来的方向：相对还原论

物理理论演变的伟大经验是，新模式涵盖的范围比其前辈理论更广泛。毕竟，牛顿定律在大多数可观测情形下完美地描述了引力。这些定律可以从广义相对论推演出来，而广义相对论效应却很难在我们的太阳系里看到——只有水星在近日点微小的“异常”运动是个例外，这时太阳的引力场非常强大。量子场论也同样如此。我们不妨将量子场想象成宇宙电脑屏幕上的像素，就像我们看屏幕时忽略像素一样，量子场论给出的是宏观距离尺度上看到的宏观物体（树木、奶牛和汽车等）的正确图像。

一种日益清晰的认识是，我们看到的物质和能量越是服从相对论，新出现的基本粒子就越多。但这些粒子确是基本的、不可分割的实体吗？坚持我们偏爱的还原论立场——即相信这些新粒子是不可分割的——就是坚持物质和时空的内在分离性。而暗能量告诉我们的是，我们必须正视这种偏见——我们必须将真空看成是相对的而不是绝对的实体；我们需要将真空的量子概念相对化。

大自然给了我们一个提示。暗能量在引力效应占主导的大尺度上才起作用，但矛盾的是，根据量子场论，它在亚原子尺度上应当接近于无限大，而在此情形下引力似乎不起任何作用。正如我已经指出的，在越来越小的距离尺度上，我们发现有越来越多的新的基本粒子。但在宏观

尺度上，量子效应可能正相反——就是说，基本粒子（譬如说电子）能够集体表现得像一个新的实体，我们把这叫作“突现的量子现象”。超导电性和超流性质就是这样的例子，在这些突现的物理现象里，集体行为是最本质的。

按照爱因斯坦告诉我们的思路，宇宙真空应当对任何观察者的运动状态保持一致，同时还展示出我们观察到的广义相对论的质量—能量和时空曲率关系。然而，观察者自身也是个问题，这个问题是由于这种虚空空间而生出的。关于观察者的定义在量子场论中和在广义相对论中是不同的。物理定律的这两种图景——突现论与还原论——之间始终存在着一种无言的紧张关系。试想如果观察者，譬如我们前面说的电梯里的乘客，只是新物理学的一个不同的侧面，那么暗能量问题怎么可能消失？

按照这种观点，基本粒子的概念——或观察者概念——就不是绝对的。也许我们在亚原子尺度上看到的所谓基本粒子实际上是空间结构本身的一种量子突现。在一个观察者看来是基本粒子的东西在另一个观察者看来却是一种突现现象。根据这种观点，一种新的相对性原理不仅将运动看成是相对的，而且基本粒子（及其相关的真空）的概念本身也是相对的。这样做的物理结果是，物质可以产生空间，空间弯曲本身也可以产生物质。我们在亚原子尺度上预期的几乎是无限大的暗能量实际上可能是大尺度上突现的时空物质的产物。这种可能性甚至能够最终解释另一个观察到的有关星系的神秘现象，在过去的半个世纪里它一直让天文学家迷惑不解：必定存在看不见的“暗”物质才能保持星系不分崩离析。

第6章
青少年中社会脑的发展

◎莎拉—杰恩·布雷克摩尔

莎拉-杰恩·布雷克摩尔（Sarah-Jayne Blakemore）

曾获牛津大学实验心理学荣誉学士学位，2000年获伦敦大学学院博士学位。目前她是伦敦大学学院认知神经科学研究所英国皇家学会大学研究员。

布雷克摩尔的研究以社会认知神经科学为中心。在伦敦大学学院，她的小组采用各种行为研究方法和影像学方法来研究青春期的智力发展、行为理解和执行功能等。他们研究的第二个重点是自闭症患者的社会认知缺陷。布雷克摩尔是《学习脑：教育课程》（Uta第五版）一书的合著者。

几十年前，很多人很难相信大脑在幼儿期之后会发生深刻变化。有些人甚至认为，我们的大脑在3岁时就基本固定了。现在，利用当代脑成像技术，科学家们在最近的研究中发现，人类的大脑在童年之后的很长一段时间里确实还在变化。某些大脑区域——特别是额叶前皮质（PFC）——在青春期甚至之后仍然在发展。额叶前皮质与大范围的认知能力——包括计划和决策——有关。它也是社会脑的一部分，所谓社会脑是指与理解他人有关的脑区网络。

自从20世纪50～60年代以来，动物（啮齿类、猫科和灵长类）实验使我们对大脑的早期发育有了很多了解。一个主要的发育过程影响到神经元或脑细胞的“线路”——即构成突触之间连接的复杂网络。在大脑发育的早期，脑开始经历突触发生，即新突触的形成。显然，婴儿大脑的这种连接数目大大超过成人水平。这之后是一个突触淘汰期——或称修剪期——这时过剩的连接会消亡。

动物研究表明，突触修剪会受到动物经历的特定环境的影响。经常使用的突触得到加强，而不经常使用的突触被淘汰。突触的修剪很像玫瑰的修剪：剪去弱枝可以使剩下的分枝长得更强壮。差不多同一时期的研究还表明，生命早期的脑发育有一个关键期，在此期间动物需要接触到大脑正常发育所需的感官刺激。早期的动物实验表明，这些过程大部分都发生在3岁左右，至少猴子大脑感觉区域的发育是这样。

在此研究的基础上，教科书常常认为，人类大脑发育的关键阶段出现在生命的前3年，因此，在此期间应给予儿童各种学习经历。但是，这种说法忽略了这样一个事实：猴子的性成熟年龄在3岁，此后不再经历人类所具有的扩展了的发育期。事实上，20世纪70年代和80年代所进行的研究表明，人类的突触生长期和突触修剪期与猴子的不同。

在20世纪70年代，芝加哥大学的彼得·胡滕洛赫尔（Peter Huttenlocher）收集了各年龄段人死后的大脑，结果发现，儿童的额叶皮质与青少年的明显不同。而脑感觉区的突触数量在童年中期达到成熟水平，额叶前皮质的突触数量则继续增加，然后在青春期开始下降。

胡滕洛赫尔的发现已在最近的利用非侵入脑成像（譬如磁共振MRI）技术对活的人体对象的研究中得到证实。这些研究表明，皮质的一些区域，特别是额叶前皮质，将在生命的前二三十年经历重大变化。[7]

利用磁共振扫描技术，人们已检测到两个主要的发育变化。首先，包括额叶前皮质在内的某些大脑区域的白质体积增大，这种增大过程会持续整个青春期直到20岁以后。白质是由轴突——一种将一个神经元的电信号携带到另一个神经元的长纤维——组成的。轴突看上去发白是因为它们覆有一层称为髓磷脂的白色脂肪物质，这种物质起着绝缘并加快轴突信号传输速度的作用。这意味着，轴突持续积累髓磷脂会长达几十年，从而加快大脑有关区域处理信号的速度。

人类大脑在青春期的第二个变化是额叶前皮质中灰质体积的减小。灰质是由神经元胞体和与其他神经元形成突触的触须样的连接体组成的。灰质发育的模式被理解为特定大脑区域中突触数量的变化的反映。

对青少年的影响

大脑的持续发育对青少年有什么影响呢？青春期是生命中以变化为特征的阶段——这是一个生理、心理和社会角色从童年向成年过渡的时期。在青春期开始时，随着青春期的到来，激素水平发生急剧变化，生理相貌也相应地改变。生命的这一时期还有一个特点，就是与情绪、自我意识、认同感和与他人关系等有关的心理变化。但最近的神经科学的研究表明，刚才所说的激素本身并不能说明这些心理上的变化。额叶皮质的发育对社会认知的影响是什么呢？

心理成熟

理解他人包括根据对方的心理状态（意图和愿望）来理解他们的行为，这个过程被称为心理成熟。脑成像实验和对脑损伤患者的研究表明，心理成熟依赖于所谓社会脑这个区域的网络。[8]

心理成熟的早期发育方面已有丰富的文献。社会能力的迹象——譬如面部识别和对他人情感的认同——早在婴儿期就已发育，这些早期的社会能力先于全面的心理成熟，后者是指能识别他人的信念，包括他们的错误信念。在典型的错误信念研究中，孩子被要求回答“萨莉”在哪里能找到她放在某个地方但在她不在房间时被“安妮”挪走了的玩具。理解萨莉不可能知道玩具的新位置的能力依赖于受试者区分萨莉的错误观念和现实的能力，大多数4～5岁的儿童都能通过这一测试。[9]

虽然心理成熟的早期发育已得到深入研究，但对儿童期之后阶段的社会认知能力发展的实证研究却少得可怜，这也许是因为大多数4～5岁的孩子甚至可以通过相当复杂的心理成熟测试的缘故。然而，作为心理成熟基础的大脑结构（包括内侧额叶前皮质）却在儿童期之后阶段经历了充分的发育。

最近的一些脑成像研究着眼于青春期社会脑的功能发育。有迹象表明，从青春期萌发到成人期开始的这段时期里，尽管心理成熟的负担在减小，但内侧额叶前皮质仍有活动。例如，最近的功能性磁共振成像（fMRI）研究运用反语理解力测试调查了人际沟通意图的发展。该实验的前提是，正如了解别人的心理状态需要具有将现实与信念分开的能力一样，理解反语要求能够将文句的字面意思与其实质意义分开。实验扫描了年龄在23～33岁的12位成年人和年龄在9～14岁的12个孩子的内

侧额叶前皮质。结果发现，儿童的内侧额叶前皮质要比成年人的更多地参与完成这项任务。文章作者解释说，儿童内侧额叶前皮质活动的加剧反映了他们需要将一些线索整合起来来解决讽刺语句中字面意思与实质意义之间的不一致。[10]

在最近的关于思考自己意图的功能磁共振成像研究中，儿童的内侧额叶前皮质的类似区域的激活强度较成年人的相应区域为高。思考按你自己的（或他人的）意图行事需要心理成熟。实验在由19个少女（年龄范围：12～18岁）组成的少女组和由（22～38岁）成年女性组成的成人组之间展开，她们被要求就设定的意图和行动方案（例如，“你想了解剧场在演什么戏。你在看报纸吗？”）进行回答。在思考这些意图时，青少年的内侧额叶前皮质比成人的更活跃。这一区域的活动随着年龄递增而下降，这表明在思考意图等问题时，青少年的神经战略与成年人的是不一样的。[11]

在青春期可能会发生内侧额叶前皮质活动的降低，因为在青春期阶段额叶前皮质要受到突触修剪的微调，因此活动量的减少对于完成这一工作是必要的。另一种（或其他）解释是，心理成熟过程中认知战略会有变化，这种变化会导致对大脑其他区域的征用。青春期内侧额叶前皮质活动的降低，到底是由于新的心理成熟战略引起还是因为修剪突触联系引起还有待观察。

对社会的影响

这种研究对社会的影响是什么呢？

首先，它质疑了这样一种观念：为了优化大脑发育，对婴儿给予额外刺激和早期正规教育是必需的。这项研究表明，人类的大脑的可塑性可长达几十年。

第二，青春期脑的明显转型表明，仅靠激素变化并不能说明典型的青少年的行为。

第三，青春期发生的突触重组很可能受到环境的影响，正如早期的突触修剪那样。目前这还是纯粹的猜测，但如果属实的话，就可以由此推断青少年有过什么样的经历。伦敦大学国王学院精神病学研究所的罗

宾·默利小组最近在《英国精神病学杂志》的研究报告暗示，经常吸食大麻的青少年比那些不吸食大麻的青少年更有可能在成年初期罹患精神分裂症。一种可能性是大麻影响青少年时期的大脑发育。

此外，青春期社会认知的发育还有可能存在性别差异。大脑发育是如何与激素变化相互作用的，以及这种互动如何影响社会认知，这些都尚未得到评估。青少年的脑发育问题是神经科学中一个新的迅速发展的领域，我们还只是刚刚开始理解。

第7章
观察大脑的互动

◎贾森·米切尔

贾森·米切尔（Jason P.Mitchell）

哈佛大学的社会认知和情感神经科学实验室主要研究员，在那里他运用功能神经影像（fMRI）技术和行为方法研究知觉是如何推断他人的思想、感情和观点的。1997年米切尔获得耶鲁大学学士学位，之后获硕士学位，2003年获得哈佛大学博士学位。他目前是哈佛大学心理学系助理教授。他也一直是达特茅斯学院和哥伦比亚大学的客座教授。他还是《社会认知和情感神经科学》杂志的顾问编辑和《神经影像与皮质》杂志的副主编。

对于不了解智人的人来说，智人似乎肯定是最不可能主宰地球的物种。与大多数其他陆生哺乳动物相比，人类显得过分瘦弱。我们没有锐利的爪子和獠牙，我们不是特别敏捷或强壮，我们没有特殊的身体适应特征让我们能够飞行，或毒杀潜在的敌人，或用变色体表来伪装自己。对于一头饥饿的狮子来说，这种无毛体弱、直立在草原上的人猿，一定是麻烦最少的午后点心。

然而，尽管有这些明显的身体上的弱点，人类却是地球上无可争议的主人——至少暂时如此。通过驯化数百种动植物，我们一直与这些物种共同生活，我们对数百种其他动植物的灭绝负有责任。我们的技术在不断改造着地球本身——土地、海洋和大气。我们差不多占领了地球的各个角落，连其他生物很少甚至根本不能停留的地方都布满了人类的足迹，如南极和低地球轨道。

我们这种看似脆弱的灵长类是如何征服地球上其他物种的呢？答案当然是，自然选择让我们配备了比利爪和獠牙更可怕的器官：人类的大

脑。

大脑由100亿个神经元组成，每个神经元平均每秒钟都能与上千个其他神经元交换信息，成人的大脑是宇宙中已知的最复杂的东西，是一台生物超级计算机，其运算能力让我们有生以来见过的最先进的硅芯片计算机相形见绌。正因为我们的大脑有海量的计算能力，人类才能够超越装备有各色武器的其他生物物种。

然而，我们这颗惊人的大脑究竟是怎样为我们提供了物种优势呢？与利爪、翅膀和剧毒的獠牙不同，我们的大脑并不与我们周围的环境有直接的互动。（如果你突然发现你的头脑与周围物体发生了接触，那么可以放心地说，事态已经非常严重了。）相反，我们的大脑是生物界里最尊贵的主角，处在半英寸厚的骨质铠甲之后，贪婪地吞噬着超过人体20%的能量。既然大脑不直接影响到它周围世界的变化，它是如何让我们在适应性方面超越其他生物的呢？

与大多数其他适应性不同，大脑需要一个翻译系统，其作用就是将它唯一的生理功能——神经元的电化学发射能力——转变成身体动作，例如语音和工具制作。我们就是靠这些动作来支配我们周围的环境的。在21世纪里，我们大多数人都熟悉类似的翻译系统：将计算机处理器的物理动作（二进制二极管的电开关）变换成计算机其他物理动作（如将显示器屏幕上的色点排列成有意义的文字或图片）的操作系统和其他软件。人类大脑使用的是一种特殊的翻译系统，叫作思想，它可以定义为这样一种算法：一组身体动作由大脑映射为另一组不同的身体动作。

描述这些算法的科学分支叫心理学，其最终目标是要将所有子程序和其他处理步骤整合起来。人就是靠这些处理步骤将来自环境的物理信息（可见光的光子，空气的振动波等）转化为其头脑内的生理表示（神经信号），然后再将它转换成对环境的实际行动（动作、语言和其他可观察行为）。

例如，视觉。人类的感知系统包括一整套处理步骤：以某种方式将光子留在视网膜上的两维图像转变成对我们周围存在的、三维空间里多彩的、有质感的物体的神经体验。要想切实意识到搞清楚这些基本功能有多么困难，是多么巨大的计算上的挑战，还真得花上一些时间。有个说烂了的（可能是杜撰的）故事，说的是马尔文·明斯基（人工智能领

域的奠基人之一）被指派来给夏季讲习班的学生讲授视觉计算问题。30年后，我们仍然在努力确认这种引导基本知觉和运动系统的翻译算法，以及人类精神舞台上出现的这些独特行为，例如，语言、推理、形形色色的社会思想。

直到大约20年前，我们关于大脑如何产生意识的大部分知识还依赖于对因中风或头部受伤而造成特定脑损伤的病人的研究，或来自对非人类动物实验的初步推断。然而，最近开发的功能成像技术和虚拟损伤技术终于使研究人员可以在大脑处理信息的同时来研究活动着的、健康人的大脑。这些技术包括功能性磁共振成像（fMRI）方法，它的前身——正电子发射断层扫描（PET）技术以及新近的近红外光谱学（NIRS）技术等。现在，当受试者执行特定任务时，这些方法已经可以让研究人员准确指向具体部位，看到这些部位代谢活动的增强（例如血流加快）。此外，更奇特的技术——譬如像经颅磁刺激（TMS）技术，即用迅变磁场产生皮质功能上的短暂电扰动（虚拟大脑损伤）——使得研究人员可以对特定大脑区域的功能暂时失去的情形下人能够做什么和不能做什么的行为进行分类。这些技术已大大加快了我们对人类意识活动的科学理解，允许我们直接看清精神活动所基于的底层硬件。

关于脑的两个假设使得这些新颖的脑成像方法可以被用于人类认知方面的研究。首先，我们一般假定不同的脑部区域有不同的处理程序组合——也就是说，不同的大脑区域可能运行着明显不同的意识子程序。其次，大脑区域被认为只能进行有限次数的运算处理，也许只是一种特定类型的计算。这两个有效的假设允许我们有力地看穿意识的组织结构。

例如，研究人员感兴趣的也许是这样的问题：我们感知和识别他人的脸（譬如“这是约翰·屈伏塔！”）是否与感知和识别无生命物体（譬如“他穿的是白色休闲服吗？”）时依赖于相同类型的信息处理机制。事实证明，这两种心智能力分属于相邻但却不同的大脑区域。这表明，不同的认知过程——输入信息的不同转换——发生在光子作用于人的视网膜的时间点与人感知人脸或无生命物体后形成意识经验的时间点之间。另一方面，如果人对人脸或无生命物体的知觉依赖于同一大脑区域的运作的话，那么研究人员则可能由此推断出，介于视网膜和人对看到的东西的理解之间的信息处理属于同类性质。如果这两个假设被证明都不正确——例如，假定不同的脑区经常执行着相同的心理活动——那么大脑

成像技术将向我们证明，心理学研究的意义极为有限。幸运的是，迄今为止，心理学家有理由相信，大脑像其他机器构造一样，是由完成专门的、各不相同的功能的各个部分组成的。

正是通过这种方式，神经影像学揭示出好些意想不到的人类心理特征——这些特征以前都未曾得到过足够的重视。例如，我们回忆过去的事情，就似乎不是一种单一的整体性操作，而是可以分为几种不同的处理机制。大脑对事情的记忆牵涉到哪些区域取决于被记住的是什么信息（面部、文字、位置等），你是需要回忆事情的具体细节还是仅仅是它的概貌，以及这种记忆编码搁置了多久。另一方面，神经影像学研究表明，一些我们认为彼此不同的心理体验实际上却依赖于相同的信息处理回路。例如，凭空想象一个对象（譬如猫的耳朵此刻是竖着还是耷拉着？）时所涉及的视觉处理区域就与你对现实世界的感知（即实际看到你面前的猫）时涉及的区域相同。神经影像学显示，某些我们感到属于单一的心理体验实则包括多个处理过程，而另一些感觉到明显有别的心理体验则依赖于相同的处理区域，这迫使心理学家必须对建立在直觉基础上的现有理论的许多方面和理论结构进行重新考虑，这些理论都是在我们能够如此轻易地检验意识活动的神经硬件基础之前建立的。

脑成像技术对心理科学的最出乎人们预料的贡献，也许是让我们对社会性思维在人类精神活动中的中心地位问题突然变得重视起来。虽然就我们具有无与伦比的、灵活的、新颖的推理能力这一点而言，主要应归功于智人阶段的成功进化，但是人类这个物种的出现，其最根本的进步却不在于个体的意识能力，而在于掌控众多个体能力的能力。事实上，人类行为的广度和复杂性——包括我们对当代世界的主宰——主要是靠这样一些机制来支撑的：这些机制使我们能够将人群组织成相互协作的各种组织，来实现个人不可能完成的目标。设计、建造并驾驶一架飞机，砌一幢楼房或一个国家的政府运作，无不需要众多的人参与。

为了将众多个人的行为拧成一股绳儿，人的意识至少必须具有两种特殊的处理能力。第一，希望能够与他人的想法进行协调的能力。我们必须有办法了解他们心中想的是什么——也就是说，要有这么一套程序，我们能够用它推断出周围的人在想什么、感受什么；他们的目标、欲望和偏好可能是什么，以及他们区别于其他人的个性特征和气质。换言之，我们必须成为“读心者”，能够觉察周围的人的心理状态。第二，我们必须具备这样的能力，即不仅能够被动地推测出他人所想的内容，

而且能够积极地影响他人的想法和感觉。了解其他人的心里所思所想的最可靠的方法是将自己的想法和感受植入他人心中。而人类就具有这种独特而强大的本领。人类的语言就可以看成是将自己的心理状态传递给他人的主要工具。人类是唯一可以直接通过指示来明确影响他人思想的动物。虽然其他灵长类动物可能也会试图操纵同类的精神状态（例如，通过欺骗），但它们似乎没有考虑过要在其他同伴的头脑中重现自己的心理状态。

社会神经科学这一新兴领域的出现表明，这些人际交往能力为人脑赋予了一些意想不到的特点。自然选择设计了专门用于理解他人思想内容的专门区域。脑成像方面的几十项研究调查了大脑在从事需要考虑他人精神状态的工作（譬如“照片上的人看上去是不是很幸福？”）时的神经活动模式与从事非心理性刺激的工作（譬如“这个人的脸是不是对称？”）时神经活动模式之间的区别。这些研究发现，大脑中存在一系列偏好“读心”的区域：内侧额叶前皮质（位于前额紧后面，与鼻子成一线的区域）、颞顶交界区域（因其位于顶叶和颞皮质交界处而得名，位置在耳后和耳上方各约两英寸的地方）和内侧顶皮质（头盖骨的紧下方）。事实上，对这些区域在人脑进行社会性思维时的观察结果可能是认知神经科学领域里最为一致的重要成果。

更有趣的是，这些脑区有着非比寻常的特点。当受试者静静地躺在MRI或PET扫描仪下不做任何具体思考时，他们的大部分脑部活动明显减少。但对脑区活动的识别发现，读心活动仍在继续。“社会脑”区域的这种持续性活动表明，人的大脑有一种琢磨别人思想的偏好。我们具有万事人格化的倾向，即将无生命物体或自然力设想成某种并不真正存在的思想的结果。这种倾向可能是这些大脑区域的社会性思考长期过度活跃造成的。人类的大脑似乎时刻准备着应付他人的思想，这种倾向使得我们感觉到这个世界充满了其他精神主体。

这些大脑区域的另一个不寻常的特点进一步强化了社会性思维神经的特殊地位：当一个人对精神以外的东西进行思考时，这些脑区往往停用。从事其他心理活动的脑区，当无需进行相应的特定信息处理时，一般来说看不出这种低水平的活动。例如，参与算术计算的脑区在你思考数字以外的东西时并不“关闭”，它们只是停止对超出其静止代谢率的刺激做出响应。但社会性思维的脑区却不是这样，它们在不使用时通常显示一种低水平的活动。社会脑区的这一特点——我们对此还知之甚少

——表明，社会性思维可能是一种与其他类型信息处理不兼容的活动。也许我们的心理算法无法对存在的其他思想和非思维性作用对象（如数学运算或无生命的物体）同时保持警觉，但一旦出现非思维性任务时，则必须暂停以社会思维方式看待世界的倾向。试想如果我们无法抑制这种将一切事物都看成是有思想的倾向，那么我们在将滚烫的水倒入大杯子，或将钉子钉入墙壁，或大力扣篮时，我们的心该有多冷酷。

最后，社会神经科学已开始显示，我们的心灵对周围不正常的思想会表现得极度敏感。这表明，我们的大脑会自发地映射周围其他大脑的活动模式。当我们看到一个人流露出恐惧的神色，我们的大脑会作出同样的反应，就像是我们自己经历着恐惧一样——这是通过激活大脑中一个叫作杏仁核的小区域产生的反应。当我们看到有人的手指被门夹了一下或被注射器扎了一针，我们的大脑反应就像我们自己正在经历着这番痛苦——这是因为视觉刺激触动了前扣带皮质。当人们看到或想到其他人遭受的痛苦时，他们经常会因这种虚拟的疼痛而龇牙咧嘴。如果我们看到另一个人闻到浓重的垃圾腐烂的味道，我们的大脑会通过激活岛叶作出恶心的反应，就像我们自己被熏到了一样。如果我们看到其他人在努力实现某个特定目标（譬如拾起某个东西），我们的大脑会通过激活位于顶叶和额叶的所谓镜区来做出反应，就像我们自己在做同样的动作。这些观察表明，人的心灵有一种试图从事与邻近的大脑进行同样的信息处理的自然倾向——我们的大脑喜欢寄存周围的大脑。虽然这个新生的发现的影响尚未得到充分的探讨，但我们乐于登录他人心理网页的冲动暗示着人类社会交往的复杂性——我们的大脑同时做着两件事情：努力找出他人的所思所想并采取同样的思维。

对大脑的社会倾向性的这些观察再次将大量的心理学研究集中到人的思维到底是如何与他人的思维合拍的问题上来。要想成功地与另一人的思维互动，或者成功地预言或影响另一个人的思维，需要一套非比寻常的认知技能。什么样的信息翻译过程能够将同伴的抬眉或一瞥转换为对他的想法和感受的理解？什么样的过程能够将一个人的心理状态转化为复杂的言论，并用听众能够理解的实际陈述加以完成（例如，我们对孩子说话就与对成人说话的方式不同）？对这些问题的答案，我们才刚刚理出点头绪，但借助于活人脑成像的新技术以及对社会性思维重要性的新认识，心理科学家可能很快就能解开我们对他人所思所想做出回应时，或影响他人或受到他人影响时思维的复杂过程。

第8章
是什么使得大观念如此牢固？

◎马修·利伯曼

马修·利伯曼（Matthew D.Lieberman）

加州大学洛杉矶分校心理学副教授，1999年获得哈佛大学心理学博士学位。他的研究兴趣主要在社会认知神经科学领域，包括自我控制、自我意识、自动化、社会性抵制和说服等。他在许多期刊上发表过文章，包括《科学》《自然神经科学》《心理学年报》《美国心理学家》《认知神经科学杂志》和《人格与社会心理学杂志》等。他的研究得到了美国国家精神卫生研究所、美国国家科学基金会、古根海姆基金会和美国国防部高级研究计划局等的资助。他的工作还得到了《时代》《科学美国人》《发现者》以及英国广播公司多媒体纪录片等方面的鼓励。利伯曼是《社会认知和情感神经科学》杂志的创始编辑，2007年，他获得了美国心理学会颁发的心理学早期职业贡献杰出科学奖。目前他正在写一本书，书名暂定为Experience Shrugged: The Rise of Simulated Experience in Mental Life and the Modern World（“经验之无奈：心理生活中模拟经验的崛起与现代世界”）。

1641年，勒内·笛卡儿出版了他的《第一哲学沉思录》。在这本书里，他介绍了他的心身二元理论，后来被简单概括为笛卡儿二元论。根据笛卡儿的理论，心灵是由明显有别于物理客体和所有物理过程的非物质灵魂来激活的。既存在着精神实体，也存在物质实体，这两者间从不见面（也许除了通过松果腺，或是通过上帝的介入，否则我们难以解释为什么心灵有了打开门的愿望，同时身体就会去执行这一预期操作这两者之间近乎完美的相关关系）。数十年后，J.J.贝歇尔出版了《基础物理学》（1667），它同样强调一种无形的实体。贝歇尔提出，所有可燃材料之所以易燃是因为它们含有燃素，一种没有颜色、气味、味道或重量的假想物质，因此火也可以用看似无形的物质激活。笛卡儿和贝歇尔

的思想曾受到广泛讨论，在他们那个时代被信奉为真理。

时代已经发生变化，因此这两种理论的命运也发生了改变。心身二元论作为上个千年中最根深蒂固的思想之一，贯穿于关于克隆、堕胎和安乐死的伦理学以及关于实验室用动物做试验的政策讨论中；而燃素说则只是在科学界被偶尔提及，而且是作为非科学理论的一个警世故事来述说的。当然，人们可能会以为，笛卡儿的二元论之所以得到保留而燃素说失宠，是因为前者已经获得了科学的支持，而后者则受到了科学的反驳。人们可以这么认为，但这种认识可能是错误的。

在科学界，这两种理论都不值得尊崇，虽然科学家仍然用二元的语言定期报告他们的研究结果。现代科学关于心灵有一条基本原则：心灵活动完全可由生物学来解释，因此属于物质实体范畴。此外，哲学上早已确立，如果离开众多错综复杂的假设，心身二元论在逻辑上是不可能成立的。但是，芸芸众生对笛卡儿描述的这种简单却难以置信的二元论抱有根深蒂固的信念。这只须看看世界各地有关心脑研究和身心研究的各种机构如雨后春笋般出现就可以明白。所有这些机构都号称要探讨这些实体之间的联系。这些机构通过指出心和物是截然不同的，需要在两者间建立联系来不断地使这种二元论具体化。目前正在讨论的有关大脑状态是如何引起心理状态，以及沉思冥想是如何用心灵来改变大脑和身体等问题同样强调了心身的区别。

在科学和哲学都已不信的情形下，为什么心身二元论思想还能这么持久地存在着？为什么关于这个问题的任何想法（或理念）能够抓住这么多的人并存在了几十年甚至几百年？一种观念如何才能成为大观念？心理学家懂得，一种信息的来源和内容是如何导致一个人拒绝或被论证说服。马尔科姆·格拉德威尔（Malcolm Gladwell）的2000年畅销书《引爆流行》（*The Tipping Point*）[12]对各种人群之所以能够形成观念传播链并确保这种概念的广泛影响给予了令人信服的通俗解释。大多数这类文化基因，或者说流行性文化观念，通常从出现到消退也就几年、几个月甚至数天。迪斯科和喇叭裤在20世纪70年代可能很酷，人们有千百种理由，条条都具说服力。但到了80年代，迪斯科和喇叭裤就过时了，街头流行的是新潮音乐和紧身牛仔裤。

但为什么像笛卡儿二元论这样的观念能够真正持久地流传下来？我认为，在于大观念往往与人类大脑的构造和功能相匹配，大脑使我们以

一种几乎不能不信的方式来看世界。我把这种解释称为迪肯理论，以纪念伯克利人类学家和神经学家特伦斯·迪肯，是他最先启发了这一解释。

在《运用符号的物种：语言和脑的共同演化》（*The Symbolic Species*：*The Co-Evolution of Language and the Brain*，1998）一书中，迪肯为人类运用的语言为什么会变成现在这种形式这一问题提供了一种反直觉的解释。迪肯指出，对语言的运用，广泛认可的解释是人类大脑的进化使然。大脑进化到一定水平，就要求人类必须使用某种语言来表达各种心理活动。迪肯却将这种逻辑颠倒过来，指出虽然人脑确实演化出能够进行符号处理的能力，但这不能说明语言本身的演化。相反，迪肯认为，正是有了语言，夫妇间才可以形成一种得到部落尊重的性信任，使得男人去狩猎时不必带上配偶。这是迪肯论点的下一部分，它对于迪肯理论至关重要。迪肯认为，语言之所以演化（并且还在不断演化），就是为了与人脑的结构和功能相适应，而不是相反。他提供的广泛的证据表明，语言的演化要比脑的演化更加快速，也更容易；随着语言在一代又一代人的交替中改变，它几乎总是变得使2岁的儿童更容易学习。

因此我们可以这么来陈述迪肯理论：大观念之所以具有持久的影响力，是因为它们适合人类大脑的结构和功能。或者如迪肯所说的那样，各种思想是按适应脑结构和功能来演变的，一种思想越是适于大脑思考，它就会变得“越顽固”。在迪肯理论适用的场合应当呈现两种效果：第一，在大观念的内容与脑结构和功能之间应该有某种形式的强有力的匹配；第二，大观念应随着时间的推移变化以更好地趋近脑组织的关键特征。本文将考虑适用于迪肯理论的两大观念：心身二元论和东西方文化。

心身二元论

虽然科学的共识是心灵和身体都是由相同材料形成的，但关于大脑是如何使日常生活中心灵和身体行为被理解的科学还处于起步阶段。事实上，关于这一问题的几乎所有证据都集中在两个方面——理解自身和理解他人。这些证据只是偶然提供了一幅大脑是如何产生其自身的心身二元论的图像。有十几项研究（主要是利用功能性磁共振成像技术）发

现，大脑的两个内侧（或中间）表面区域，一个位于额叶前皮质（称为内侧PFC），另一个位于顶叶皮质（称为内侧PAC），在内省（即一个人专注于自我，反观自己的状态、特征或偏好）时更容易活跃。[13] 另一项研究调查了有关识别自己身体体征（如通过视觉辨认自己的脸）的脑区。出人意料的是，在神经成像研究中，当向受试者显示他们的面孔照片时，内侧PFC和内侧PAC这两个主要涉及非身体属性的区域并不激活，而位于大脑外表面的外侧PFC和外侧PAC区域则被激活。[14] 此外，内侧PAC似乎还与观察自己的身体运动有关，对这一区域实施干扰会产生灵魂出窍的体验，在某种意义上——如精神分裂症——好像是有人正在控制自己的身体。[15]

类似的区别在研究他人的脑处理过程中也可以看到。具体情形取决于测试科目是要理解他人的意识活动还是要理解他的身体动作。所谓“心理化”实验，就是让受试者琢磨他人心里在想什么——他的意图、信念或感受。与心理化直接相关的大脑区域是内侧PFC区域。这个心理化区域靠近（但不全同于）自我反省所涉及的区域。因此，对他人的思想或自己的思想进行心理化测试属于内侧PFC研究的新宠。

当我们看明白他人的身体动作但却无意弄懂他心里在想什么，这时会怎样呢？例如，我们模仿他人轻敲手指的动作但不必考虑这个人的心境。在这种情况下，观察表明，外侧PFC和外侧PAC区域的活动是一致的。这些区域通常合称为镜像神经系统，因为对其他灵长类的单元记录（single-cell recordings）表明，灵长类动物是执行一个操作（例如够取食品）还是只看着其他同类做这个动作，其外侧PFC和外侧PAC区域的神经有着同样的反应。[16] 如同自我处理（指自省或识别自己体征——译注）时的情形，我们看到，在琢磨他人时，我们关注的是他们的意识活动还是他们的身体动作，这之间我们的大脑活动有着明显的区别。

当我们试图理解目标的意识活动时，不论这个目标是自身还是他人，我们大脑的内侧活动是主要的；当我们将注意力集中在目标的身体活动方面时，外侧活动则占支配地位。这些脑区都处于大脑表面内、外侧的相对类似的位置（即PAC和PFC）上，但它们却十分的不同，一个关注的是心灵，另一个关注的是身体。此外，外侧区域的活动与内侧区域较低水平的活动相关联，[17] 这表明——至少在某些情况下——内、外侧活动可能存在竞争性。

因此在大脑内，心和身由不同的网络代表，由此产生了一种脑内二元论。一般来说，当大脑通过不同的脑网络处理两件事情时，这两件事是按不同类别来感受的。例如，颜色和数字就是按单独的类别来体验并在分开的神经网络里分别处理。[有趣的是，有极少数具有联觉的人能够从数字上看出颜色来，或将其他“感受性”混在一起——例如，“看”音乐或“品尝”视觉刺激。加州大学圣迭戈分校的神经学家拉马钱德兰（V.S.Ramachandran）证明，这些人往往是用同一脑区来处理这些不同的感受性。]由于大脑有这种正常分离，因此试图说服人们将精神和身体看成是一种东西而不是两种东西，就好比要说服他们将颜色和数字说成是一种东西。科学告诉我们什么都可以，只是不要与我们眼前的日常经验相矛盾。

回想一下，迪肯理论的第二个要点是说：牢固的观念都是由不牢固的观念经过观念转变以便更好地适应脑结构和功能演变而来的。心身二元论似乎就出现了这种观念的演变。当笛卡儿阐述这一观点时，二元论几乎算不上是一个新的观点。在他之前提倡这一观点的就有毕达哥拉斯、西塞罗、圣·奥古斯丁和托马斯·阿奎那。前笛卡儿二元论的最著名人物当属柏拉图，他提出了一种与物理世界形成对比的普遍形式的世界理论。譬如说，我们之所以能够将一把具体的椅子看作各类椅子中的一员，是因为我们具有“椅子”这一普遍的理念。这些普遍理念存在于它们自己的王国里，而不是存在于我们心里或是体内。虽然柏拉图的理论在哲学圈子里有着广泛的影响力，但它从来没有成为大众的共同理念，也从来就没有一项社会政策能以普遍形式反映出我们的感受。难道这是因为大脑不具备用于处理普遍形式的结构？抑或这种普遍形式只是大脑这一全能符号机器可处理但不必处理的无数个命题方案中的一种？就好像学生宿舍里“发现”我们可以将行星绕日旋转类比到电子绕原子核旋转，或将银河系看成仅仅是巨大的宇宙实体中的一个分子那么不值一提？我们可以接受地球/电子之间类比的想法，但这种想法谈不上牢固，也不具有普遍形式对物理世界的那种二元论关系。

各种各样的二元论被提出来用以解释世界上多种并存的复杂性。然而，在笛卡儿二元论出现之前，这些理论没有一种能够真正流传开来。笛卡儿版本则刚好与大脑处理心身关系的主要功能合拍。因此尽管科学家和哲学家共同努力来抵制心身二元论，但它仍作为一种核心理念和看待世界的方式留存下来。

让我们转到能够用迪肯理论来说明的第二个大观念。

东西方文化的异同

20世纪90年代初以来，心理学领域关于一种特定文化形态是否会以及如何对产生这些文化的心灵造成影响的争论日趋白热化。概念上的突破（由此已引出数百项研究）始于1991年。这一年，斯坦福大学的海泽尔·马库斯（Hazel Markus）和京都大学的忍北山（Shinobu Kitayama）提出，东方文化倾向于教育孩子从相互依存的角度看世界以及个人在其中的位置，而西方文化则倾向于灌输独立地看世界和个人在其中的位置。从本质上讲，东亚的教育是要让人相信，我们都是相互关联的，集体的需要大于个人的需要。相反，西欧和北美人学到的是自己的目标、感觉和成就是第一位的。社会的奖惩也遵循相应的规则。例如，在强调相互依存的（东方）文化里，“枪打出头鸟”；而在强调独立的（西方）文化里，“爱叫的鸟儿有食吃”。因此，在这种或那种文化里培养出来的人必定会带着他们被塑造的心灵来看世界，即用相互依存的眼光来看世界，或以独立的方式看世界。由此导致每个人按照他们文化的价值观来生活。

每一种文化的价值观代表了该文化特定的大观念，它们在各自的文化中已经存在了上千年。标准的理解是文化塑造心灵和大脑。而迪肯理论则认为反过来解释可能也成立。也就是说，如果东亚人和西欧人的脑子不是现在这样，那么每种文化的大观念还能这么牢固吗？事情是否会是这样：不同地理区域的人的大脑上的差异促进了各自的文化叙述，而这种叙述导致每个群体按其自身标准来评价社会形态的各个方面，这种社会形态反映了脑组织的群体特征？例如，如果一种文化里的人先天就听觉不灵，而另一种文化里的人先天就视力很差，那么无疑他们对音乐和艺术的评价就会不同。

鲍德温·维（Baldwin Way）是我在加州大学洛杉矶分校实验室的博士后研究员，他最近提出，东西方血统的个体之间存在一种重要的基因上的差异，这种差异对他们的大脑有着不同的影响。我们进行了一系列交谈，决定开始检验这一想法。维对控制大脑的5-羟色胺系统的基因研究进行了调研。他发现，东西方血统的个体在5-羟色胺转运基因（5-HTTLPR）在调节区内的分布变化上表现出差异。基于两个等位基因的

不同组合，5-HTTLPR的基因多态性表现为3种不同形式，（为方便起见）分别被称为短—短型，长短型和长—长型。2/3的亚洲人具有短—短型，而具有这种类型的美国人和西欧人则仅占1/5。这是一个巨大的、高度可信的差异，它已在多项研究中被观察到。[18]

5-羟色胺系统，尤其是这个5-HTTLPR基因，与社会情感的敏感性有关。例如，人们在一项研究中发现，具有短—短型基因的儿童均表现出较高的抑郁症风险，但前提是他们缺乏社会鼓励；而那些具有长—短型和长—长型基因的儿童患抑郁症的风险则不受社会鼓励因素的影响。[19] 另一项研究发现，来自非鼓励性家庭且具有短—短型基因的个人患抑郁症的风险最大，而虽具短—短型基因但来自鼓励性家庭的人患抑郁症的风险最小。具有长—短型和长—长型基因变异体的个体则处于中间，不论他们的家庭背景是鼓励性的还是非鼓励性的。[20] 这些结果表明，那些带有短短型5-HTTLPR基因的孩子的健康更多地依赖于社会环境的品质，一般来说，这些人对社会环境可能更加敏感。

根据迪肯理论，东亚裔个体多具有短—短型5-HTTLPR基因多态性这一特征暗示着，他们可能具有一种易于抱团的神经化学基础，并在此基础上建立起文化价值观，或在世界的这一地区使得这种大观念传承下来。如果你的健康很容易受到别人是怎样对待你的影响，那么你就肯定会偏爱一种鼓励他人以促进你的健康的文化。相反，西方人相对缺乏这种类型的基因，这一点导致他们更容易创造出一种看重独立和个人成就文化的神经化学基础。

回想一下，心身二元论在笛卡儿之前就已出现，然而，自从他提出了他的二元论形式体系后，这一观念却没能进一步发展，尽管众多的批评也是个原因。在东西方文化差异这个问题上，很长时间以来一直有一种将它看作是观念演化的观点，即认为这种文化差异是两种文化观念的地域性迁移的结果——两者似乎都起源于中亚，后来一个几乎完全向东迁移，另一个则向西迁移。

我们可以合理地将东西文化描述为各自宗教和特定公民理论的融合。东方文化固化为新儒学，它结合了佛教信仰，认为我们都是互相依存的，自私的品格在儒家看来是不健康的，它刻画了一种建立在社会成员之间义务关系基础上的社会。西方文化则脱胎于犹太教和基督教神学的组合，它信奉个人对自己永恒的救赎负责和希腊公民的生活态度，它

强调体己和自由意志。

印度的佛教向东亚国家传播，而且越往东传播得越持久，以至于佛教在印度不再是主流。考虑到印度只有不到一半的人口带有短短型5-HTTLPR基因，这种变化很容易理解。相反，基督教的传播则是从中东一路向西，先是传播到欧洲，然后到北美地区，而且越往西越持久。在这两种情形下，大观念开始时都是相对较小的观念，传播到数千英里之外才找到可以使之蓬勃发展的适宜地区，才逐渐发展成为大观念。有趣的是，在亚历山大大帝时代，在罗马帝国时代和后来的中世纪时期，东西部宗教领袖和代表有过多次接触，但这些宗教间相互得益的交流则几乎没有出现过，虽然宗教在自然扩张的方向上被接纳时相对较容易——例如，佛教向东扩张。这种情形直到最近，当20世纪的全球经济发展令这些限制黯然失色后，情况才有好转。这些文化的大观念似乎还将迁移，直到它们找到具有适当的神经化学基础的人口才能使它们变得牢固。

我们喜欢把我们的信念看成是某种逻辑分析和同行影响相结合的结果。迪肯理论则提供了另一条思路：人类的大脑倾向于寻找一些有吸引力的想法，因为大脑的结构与所提问题里的概念结构之间存在匹配性。对于笛卡儿二元论的情形，我们已经看到，大脑是以两套分立的神经回路来表达心和身的。很可能是我们先有了心和身分属于不同类别的直接经验，然后才有的心身二元论理念，尽管这方面的证据还充满矛盾。对于东西方文化差异的情形，我们已经看到，基因变异的区域特征产生出明显有别的大脑化学物质，从而使得两群人对社会信息反馈的敏感性存在明显差异，因此在文化信仰和价值观的接受上也存在差异。这种差异主要表现为需不需要优先考虑社会的相互依存关系。在这两种情形下，迪肯理论利用神经学说为我们提供了对我们深信的某些信念的反直观的解释。当有足够多的大脑倾向于寻找同一种有吸引力的观念时，这种观念很可能坚持相当长的时间。

第9章
道德心灵的生理学基础

◎约书亚·格林

约书亚·格林（Joshua D.Greene）

认知神经科学家和哲学家，1997年获哈佛大学学士学位，2002年获普林斯顿大学博士学位。2006年任哈佛大学心理学系助教。他的主要研究兴趣是道德的精神病学和神经科学研究，重点是道德决策过程中情感和“认知”之间的相互作用。他的更广泛的兴趣是哲学、心理学和神经科学之间的交叉。目前他正在写一本有关道德的新兴科学认识的哲学意义的书。

考虑下面的道德困境：这是在战争年代。你和你的乡亲们正在地下室躲避附近的敌人。这时你的宝宝开始哭啼，你掩其口来阻止声音扩散开来。如果您将手拿开，宝宝的大声哭泣就会被敌兵听到。他们会找到你、你的宝宝和其他人，他们会将你们全部杀死；如果你不放开你的手，宝宝就会窒息。为了挽救自己和其他村民而将婴儿捂死，这在道德上可以接受吗？

思考这些问题会令人感到不安，但却有启发。这种困境，即所谓哭婴困境，很好地反映了道德观念和政治观念这两个主要学派之间的紧张关系。一方是如边沁（Jeremy Bentham）和穆勒（John Stuart Mill）这样的功利主义哲学家，按照他们的哲学，行事是否道德最终要由能否生产最佳的整体后果来衡量，追求的是“更大的利益”。另一方是像康德（Immanuel Kant）这样的道义论哲学家，认为权利和义务往往胜过更大的利益。在哭婴困境里，更大的利益（至少就挽救生命的数目来衡量）是由扼杀弱小生命来实现的，但很多人认为，扼杀弱小生命，除了令人悲痛和难以下手之外，在道义上也是错误的——是对婴儿的权利、父母的责任的背离，或两者兼而有之。

哭婴困境也是窥视人类大脑结构的一个窗口。人们通常谈论“伦理本能”或“道德意识”，即暗示着道德判断是一个统一的现象，但对于道德判断的科学研究的最新进展却描述了一幅非常不同的景象。道德判断似乎依赖于直觉的情绪反应与需付出更多努力的“认知”过程之间复杂的相互作用。更具体地说，对有害行为的直觉的情绪反应（“不要捂死孩子！”）似乎依赖于一套大脑系统，而我们更可控的认知反应（“窒息婴儿以促进更大利益”）则依赖于不同的大脑系统。我们之所以对这种道德困境大伤脑筋，正是这些神经系统处于竞争状态，我们的痛苦感觉正是这些竞争的结果。如果我是对的，那么竞争性神经系统之间的这种紧张关系就不仅是几百年来穆勒和康德的拥护者之间争论的根本原因，而且也是当前诸如干细胞研究和折磨人的对恐怖分子疑虑等问题带来苦闷的根本原因。

让我们考虑两种道德困境，它们是当代伦理学中所谓“巡道车问题”思想实验的一部分。其中第一个，我们称之为道岔困境。它是这样的：一台失控的巡道机车将要辗死5个人，但你可以扳道岔将它移至岔道来挽救这5人，在岔道上机车将只会辗死1人。那是不是就该扳道岔呢？在此情形下，大多数人会说是，这与功利主义哲学是一致的。第二种情形是考虑人行天桥上的两难境地：这里，失控机车仍然对5个人的生命构成威胁，但此时你不是站在道岔旁，而是站在横跨轨道的天桥上，一边是迎面而来的失控机车，另一边是对危险毫不知情的5个人。你旁边站着一个高大的男人，挽救5人生命的唯一办法就是将这个大男人推下天桥阻塞机车行经的轨道，这样才能停止杀人的机车，但需要你的朋友充当横木。那是不是就该将这个人推下去以挽救那5人的生命呢？（我知道你在想什么，我不吃这一套：你不能自己跳下去，你没有大到可以让机车停止；你也不可能对轨道上的人大喊有危险，那没用，他们听不到。是的，机车肯定会杀死全部5人。你的处境只能两者择一：推还是不推？而且这个大男子确实能够阻止机车行进。这个大男子不是乌萨马·本·拉登，铁轨上的5个人也不是你的父母、你的两个孩子和私人教练。总之，你不可能重写剧本使问题变得容易一些。）在这种情形下——经过适当的解释——大多数人判断，通过牺牲1个人的生命来拯救5人是不对的。在这里，康德哲学占了上风，因为大多数人认为天桥上的人的权利要重于更大的利益。

为什么在第一种情形下我们看重多数人但在第二种情形下则否呢？几年前，我的感觉是，天桥困境中的行为，即近距离故意将人推下天

桥，要比道岔困境中的行为有更强烈的情绪体验，而正是这种情绪反应上的差异能够解释为什么我们在这两种情境中的态度是如此的不同。为了检验这一假说，我和我的同事对人们思考这两种情形时的大脑活动进行了扫描，这里我们将天桥困境称为“个人困境”，将道岔困境称为“非个人困境”。我们的假设预言，个人困境会引起与情感有关的大脑活动的增加，而非个人困境则更多地引起脑区域中与需要付出努力的认知过程（如推理）相关的活动的增加。这是我们要求证的一个方面 [21] 。更具体地说，个人困境（如天桥困境）下产生的反应将引起内侧额叶前皮质活动的增加，以及脑中其他与情绪和社会性思考相关的区域活动的增加。相反，非个人困境（如道岔困境）则引起背外侧额叶前皮质活动的增加，即脑中典型的“认知”区域活动（譬如说，你记住一个电话号码）的增加。

在道德思考方面，这能告诉我们什么呢？这里的想法是：面对道岔困境和天桥困境，人们诉诸功利性推理。（“以1抵5？听起来合算。”）但对于天桥困境设定的更加个体化的伤害，也存在负面的情绪反应，说：“不！不能将那个人推下去！”而且这种反应往往主导决定。在道岔困境中人们的情绪反应相当弱，因此功利性推理会主导决定：我们投票决定救那5人。在天桥困境中主导决定的情绪反应依赖于与情绪有关的大脑区域如内侧额叶前皮质的神经活动，而在道岔困境中主导决定的更缜密的思考则依赖于典型的大脑“认知”区域如背外侧额叶前皮质的神经活动。

在接下来的实验中，我们来考虑如哭婴困境这样的更为困难的情境。这些也都属于个人困境，但它们构建需要你有更强的功利上的理由。在天桥情形中，是以一命抵五命；但在哭婴困境中，如果你不采取行动，则每个人都将死亡，包括你和你的宝宝。面对道岔困境和天桥困境，人们的判断相当一致；但在哭婴困境下，人们的判断分裂为差不多一半对一半，几乎每个人都需要很长的时间才能做出响应。

这是怎么回事？如果我所描述的理论是正确的，那么哭婴困境一定是触发了大脑中情感区域和认知区域之间的冲突。说得浅显点，就是大脑中有一个称为前扣带皮质的区域，它能对这类内部冲突做出可靠的响应。当你的大脑试图同时进行两项不同的工作时，前扣带皮质就会说：“休斯顿，我们出问题了。”我们预计，面对哭婴困境，这个区域会变得更加活跃，事实也果真如此。

如果前扣带皮质说："休斯顿，我们出问题了。"这自然就引出了一个问题："休斯顿"在哪里？事实证明，"休斯顿"位于背外侧额叶前皮质，就是我们记住电话号码、进行抽象推理的那个地方。大脑的这个区域也赋予我们抵制冲动的能力。这些行动的共同特征是认知控制——一种引导注意、思考以及按既定目标或意图采取行动的能力。如果以"个体"方式去伤害某个人的想法触发了使我们说"不"的情绪反应，那么为追求更大利益而同意进行"个体"伤害就要求有能力去战胜情绪反应。也就是要求增强主管认知控制的背外侧额叶前皮质的活动。这表明，当人们对像哭婴情形这样的困境作出功利性的判断时，他们的背外侧额叶前皮质应当表现出增强了的活动能力，这是我们要求证的另一点。[22] 我们最近的研究结果完全切合相同的模式。我们让受试者考虑这样一种困境：为了增进更大的利益需要违反承诺，结果一如既往，我们看到，在人们给出有利于更大利益的功利性答案时，他们的背外侧额叶前皮质的活动明显增强。

道德判断的这种二过程理论——这里"二过程"是指不同的情感和认知过程——可用于对神经病病人的行为进行一些有趣的预测。例如，众所周知，额颞叶痴呆（FTD）病人呈"情感钝化"症状。加州大学洛杉矶分校的研究小组报告了FTD病人对道岔困境和天桥困境的反应。他们对道岔困境的反应是相当标准的，但他们要远远比其他人更有可能同意将男子推下天桥。由于缺少提醒他们说"不"的情感，这一行动似乎还是一个"很好的解决办法"。另外两个研究小组，一个在艾奥瓦州，另一个在意大利，在检查前内侧额叶皮质（脑中一个主管情感基决策的重要区域）受损病人时也得到了类似的结果。这两个小组发现，这些病人对天桥困境和哭婴困境给出了不同寻常的功利性反应。事实上，艾奥瓦州的病人给出的功利性反应几乎是对照组的5倍。[23]

东北大学的皮耶尔卡罗·瓦尔德索洛（Piercarlo Valdesolo）和戴维·德斯提诺（David Desteno）用聪明但技术上却并不复杂的方法得出了同样的结论。他们陈述了受试者在两种不同条件下面对道岔困境和天桥困境时的反应。在实验开始时，让一些人观看了从电影《周六夜生活》节选下来的有趣片段，而其他人则观看没有特定情感内容的电影剪辑。结果发现，人们对道岔困境的反应不受电影选择的影响，但看了有趣的《周六夜生活》剪辑的受试者在同意将男子推下天桥方面却提高到近4倍。这里的概念是，积极的情绪体验可以抵消将男子推下天桥带来的负面情绪。

我和我的同事进行了一项类似的实验，目标是认知控制过程而不是情绪体验。在我们的实验中，受试者必须在紧盯着电脑屏幕上滚动数字的同时做出判断，每次出现数字5就按按钮。这种令人心烦的任务称为“认知负荷”测试，目的是要干扰基于外侧额叶前皮质的各类高级认知过程。我们发现，认知负荷使人们在给出功利性答案（“以更大利益的名誉扼杀婴儿”）时变得缓慢，但在给出典型的道义论答案（“不要扼杀婴儿，即使每个人都会死”）时则没有这一效应。（事实上，认知负荷似乎加快了道义论答案的判断，但这种效应整体上并没有显示出统计学上的重要性。）这两项研究犹如镜像：阻碍情绪过程将使功利性判断变得更容易；阻碍受控的认知过程将使功利性判断变得更慢。

这些结果其实是哲学家会感到惊奇的一般模式的一部分。当明显的道义责任（“不要用人去阻止机车”）与更大利益（“最好去拯救那5条生命”）相冲突时，有利于责任的判断由情绪驱动，而有利于更大利益的判断则更多的是由受控的认知过程驱动。这一点令人惊奇，因为像康德这样的将责任看得重于更大利益的哲学家往往被视为理性主义者，这些哲学家的道德结论是基于理性。但这里所描述的研究却表明，这种哲学里推理的成分较少而更多的是合乎理性。我的同事乔纳森·海德特（Jonathan Haidt）认为，几乎所有的道德推理都是如此，但我不同意。[24] 基于这里描述的研究，我相信功利性判断是由推理过程来驱动的，要借助于背外侧额叶前皮质（又名休斯顿）才能奏效。当然，我不认为这一切会如此利落简单。按照大卫·休谟（David Hume）的理论，我怀疑甚至功利性计算也需要情感——与其说像发出警报，倒不如说更像沙粒在秤盘上堆积——脑成像数据的确暗示存在这种迹象。

人们时常问我为什么要费心搞出这些奇怪的假设性困境。难道我们不应该研究真正的道德决策吗？在我看来，这些困境就像一个遗传学家手中的果蝇。他们在实验室里可以得到很好的管理，但其复杂性足以牵连到越来越广阔的外部世界的一些有趣的东西。考虑到这一点，让我来介绍最后两个道德困境，最初设计它们的是功利主义哲学家彼得·辛格（Peter Singer）。[25]

一天你走过一个水池，发现有个小孩快要淹死了。你可以很容易地涉水过去救她，但是这会毁了你时尚的新款意大利西装。于是你一走了之。你这人是不是很冷酷？是的，我们说。下一个情形：您收到一封来自令人尊敬的国际援助组织（譬如联合国儿童基金会或英国牛津饥荒救

济委员会）的信。他们希望你捐出500美元，用以挽救急需食品和药品救助的贫穷非洲儿童的生命。你为这些儿童感到难过，但是你一想到要买时尚的新款意大利西装，您还是希望为此省下这笔钱。于是你将信件丢进了垃圾桶。你这人是不是很可怕？我们说，你不是圣人，但你肯定没有做错事。

拒绝挽救一个在你面前将溺毙的孩子与拒绝挽救一个在地球另一端贫困地区的孩子，这两者之间到底有什么区别？在这个问题上你的合乎理性的心灵已经在工作：面对溺水儿童，你是唯一可以援手帮助的人；但对于贫穷的非洲儿童，则可以有许多其他人去提供帮助，那是世界性问题，不是你能解决的。公平得很。但如果现在情形换成是你与其他很多人（这里正在召开美国律师协会的年度会议）站在水池周围看孩子淹死，他们也都相当喜欢他们自己的时尚的新款意大利西装，那又该如何看待呢？就任孩子淹死？我们可以将这个游戏玩上一天，但不太可能找到满意的解决办法。另一种做法是考虑有关的心理学及其自然史。

让我们用岔道困境和天桥困境来尝试第一种情形。正如我已经解释的，同样是挽救5个人的生命，推人致死需要比扳道岔付出更多的情绪体验。但这是什么原因呢？从进化上看可能是有用的。我们是在非常老套的盛行推挤的环境中进化的，而不是在一个用机械解决威胁的时代里发展起来的。因此，“个体”暴力这样的基本形式触发了我们的道义开关是有道理的，而鲜明的现代形式的暴力行为则没有这种功能。从这个意义上说，类似的利他行为及其支撑这种行为的情感可能也可以由此得到解释。我们不是在牺牲时尚去挽救遥远陌生的生命那样一种环境中发展起来的，而是在帮助此时此地绝望的人的环境中发展进化的。大自然赋予我们一副可以拉动的心弦，生物的关键特性就在于其生存取决于彼此间的合作。但是，大自然不能预见到可能有一天我们的生存将依赖于跨越海洋和大陆的合作，因此忽视了为我们装备一副可以随时从远处拉动的心弦。

当然，我们是非常聪明的物种。通过运用智慧，我们已经使自己变得比地球上所有其他生物更快、更强，也更危险。或许通过将复杂的认知能力运用于解决现代生活面临的问题，我们能够超越我们的道德本能的局限性。

第10章
语言是如何影响我们思考的？

◎莱拉·博罗迪茨基

莱拉·博罗迪茨基（Lera Boroditsky）

斯坦福大学心理学、神经科学和信号系统助理教授。博罗迪茨基博士出生于苏联的明斯克，2001年获斯坦福大学认知心理学博士学位后，曾执教于麻省理工学院脑与认知科学系，后回到斯坦福大学任教。她还管理着印度尼西亚雅加达的卫星实验室。

博罗迪茨基的研究以智力表现的性质以及知识如何从心灵、世界和语言的相互作用中产生等为中心。其中的一个重点是调查语言和文化对人类思维方式的影响。为此，博罗迪茨基的实验室在世界各地——从印度尼西亚到智利、土耳其、澳大利亚原住民——收集数据。她的研究已在媒体上广泛报道，并赢得了包括美国国家科学基金会颁发的职业成就奖和塞尔学者奖在内的多个奖项。

人类互相交流使用的语言令人眼花缭乱，每一种都极尽可能不同于另一种。那么我们的语言会影响到我们看待世界的方式、我们的思维方式以及我们的生活方式吗？操不同语言的人的思维方式不同，这种不同仅仅是因为他们操不同的语言所致？学习一种新语言会改变你的思维方式吗？会多种语言的人在说不同的语言时是不是思维方式也不同呢？

这些问题几乎涉及精神研究领域的所有重大争论。它们将一大批哲学家、人类学家、语言学家和心理学家卷了进来，它们对政治、法律和宗教有着重要影响。然而，尽管不断受到关注和争论，但对这些问题的实证研究工作却少得可怜，这一状况直到最近才有所改变。很长一段时间里，语言可能决定着思维特点的观点被认为是最不可能得到验证的，而且这种检验往往都是完全错误的。我在斯坦福大学的和麻省理工学院

的实验室的研究有助于重提这一问题。我们已在世界各地——中国、希腊、智利、印度尼西亚、俄罗斯、澳大利亚原住民——收集了数据。我们得到的结论是，操不同语言的人确实有着不同的思维方式，甚至语法细节都会深刻地影响到我们如何看待世界。语言是人类特有的礼物，是人类经验的集中体现。理解它在建设我们精神生活中的作用将使我们能够更深入地了解人性的本质。

我经常通过向学生提出以下问题来开始我的本科生讲课：你最不愿意失去的是哪一种认知功能？他们大多数人选择的是视觉；有些人选择听觉，偶尔也有个别爱说俏皮话的学生可能会选择幽默感或时尚感。几乎没有人自发地说到最不愿失去的是语言能力。然而，如果你丧失了（或先天就缺失）视力或听力，你仍然可以有一个奇妙丰富的社会存在。你可以有朋友，你可以接受教育，你可以拥有一份工作，你可以建立一个家庭。但是，如果你从来就没学过一种语言，你的生活将会是什么样子？你还能有朋友，还能接受教育，能拥有一份工作和建立家庭吗？语言对于我们的经验是如此基本，并且如此深刻地成为我们所以为人的一部分，以至于没有它我们很难想象我们的生活会是怎样的。但语言仅仅就是表达我们思想的工具，抑或它们实际上也影响着我们的思维方式。

有关语言是否和怎样决定着思维特点的大多数问题都可以从如下简单的观察开始：各种语言彼此不同，而且差异很大！让我们举个（完全是）假设性的例子。假如你想说，“布什阅读乔姆斯基的新书”。我们不妨将注意力集中在动词“读（read）”上。用英语说这个句子，你必须标出这个动词的时态，在本例中，我们必须像发“red”而不是像发“reed”那样来发这个音。但在印尼语里，你就不必（实际上，你也不能）改变动词的时态。而在俄语里，你在改变动词时则必须标出它的时态和性（阴性还是阳性），因此如果读书的是劳拉·布什而不是乔治·布什，你就必须使用不同的动词形式。在俄语里，动词还必须包含指称的动作是否完全的有关信息，因此如果乔治只是读了本书的一部分，你用的动词就不同于描述他读完整本书时所用的动词。在土耳其语里，动词里必须包含你是如何得到这些信息的：如果你是用自己的双眼目睹了这一不太可能发生的事情，你使用的是一种动词形式；但如果你只是简单地读到或听说这件事，或是从布什说的东西里推断出这件事，你会使用不同的动词形式。

显然，语言在不同人的嘴里需要不同的形式。这是否意味着发言者对世界的思考方式有所不同呢？英语、印尼语、俄语和土耳其语的发言者最后以不同方式处理、区分和记忆他们的经历，这仅仅是因为他们操着不同的语言？

在一些学者看来，这些问题的答案显而易见是肯定的。只要看看人们交谈的方式，就可以说明这一点。当然，不同语言的使用者必须处理好如何将世界惊人的不同方面准确地表述出来，只有这样，他们才能够恰当运用自己的语言。

争论另一方的学者不是去寻找人们交谈时那种言之凿凿的差别。他们认为，我们所有的语言表达只是我们可获得的信息中很少的一小部分。尽管操英语的人用动词时不能像说俄语和土耳其语的人那样包括同样多的信息，但这并不意味着操英语的人不注重同样的东西，区别仅在于他们此刻没提到这一点。这种情形不是没有可能：每个人都在以同样的方式思考，注意到同样的事情，但却操着不同的语言。

相信跨语言差异的人反驳道，对同样一件事情每个人的注意点不可能相同：如果每个人的注意点都相同的话，那么人们就会觉得学另一种语言很容易。不幸的是，学习一种新语言（尤其是与你熟悉的语言不密切相关的那种）是很不容易的，这需要注意一整套新的区别。无论是明显有别的西班牙语调、土耳其语调，还是俄罗斯语调，学习讲这些语言要比仅仅学习词汇需要掌握更多的东西：它需要关注世界上有影响的大事，这样你说的话才能富含正确的信息。

这种对语言是否影响思维的先验争论已经进行了几百年，有的人认为语言不可能决定着思维特征，另一些人则争辩说语言不可能不对思维特征产生影响。最近我的小组和其他人已经找到了实证检验这类古老的辩论中某些问题的方法，并取得了非常令人感兴趣的结果。因此，我们不争论什么必然是正确的或什么不可能是正确的，而是搞清楚什么是真实的。

我们先到澳洲北部约克角城西边的原住民社区波姆普拉（Pormpuraaw）来看看。我来到这里是因为当地的库克塔约尔人（Kuuk Thaayorre）的说话方式。库克塔约尔人在说到某物相对于观察者的位置时不是用英语里通常使用的“左”“右”“前”或“后”等字眼，而是

像其他原住民部落那样，用主方位词——北、南、东和西——来表达空间概念。[26] 他们在所有尺度上都这么用。因此譬如你得这么说“你的东南腿上有只蚂蚁”或“把杯子往西北偏北挪一点”。操这种语言的结果明显是，你要时刻留意方位，否则你无法正常说话。在库克塔约尔人那里，通常的问候语是“你上哪儿去？”而回答应当是类似于“东南偏南，距离不远”。如果你不知道你所处的方位，你甚至不能用一句“你好”就敷衍过去。

因此，主要依赖于绝对参照系的人（如库克塔约尔人）与依赖于相对参照系的人（如英国人）在航海能力和空间知识方面有着深刻的差异。[27] 简言之，与说英语的人相比，说像库克塔约尔人说的语言的人对空间方位和自身所处位置（即使是在不熟悉的地方或不熟悉的建筑物内）的敏感性要高得多。使他们能够如此的——事实上，他们不得不如此——正是他们的语言。语言上的这种训练使他们成就了一度被认为超出人的能力的航海上的壮举。

因为空间表达方式是这样一种基本的思维域，人们思考空间问题时的差别不只是局限于空间问题。人们是依靠自身的空间知识来建立其他更复杂、更抽象的表达能力的。业已证明，像时间、数量、音乐音调、亲属关系、道德和情感等的表达都依赖于我们如何进行空间思考的。因此，如果库克塔约尔人对空间概念有不同的思考方式，那么他们对时间这样的其他概念的思考是不是也与我们不同呢？这正是我和我的合作者艾丽丝·加比（Alice Gaby）来波姆普拉要寻找的东西。

为了验证这个想法，我们给当地人一组表现事物某种时间进展的图片（例如人逐渐变老、鳄鱼逐渐成熟或香蕉被吃掉过程的照片）。他们要做的就是将摆在地上的弄乱了的图片按正确的时间顺序重新整理好。我们让每一位受试者分两次坐在不同朝向的椅子上进行测试。如果你让说英语的人做这事，他们理好的图片时间顺序是从左至右，而希伯来人则往往会按从右到左的顺序排列，可见方向在语言文字的书写上起着一定的作用。[28] 而像库克塔约尔人，他们不使用“左”和“右”的字眼，那将怎么办？他们是怎么处理的？

库克塔约尔人排列图片既谈不上是从左到右多点还是从右到左多点，也谈不上按距离身体的远近，但他们的排列并不是随机的，而是有一种不同于讲英语的人的排列模式：当他们面南而坐时，他们安排的时

间顺序是从东到西，或者说，从左到右；当他们面北而坐时，图片排列则是从右到左；当他们面东而坐时，图片排列是由远及近，等等。即使我们没有告诉受试者任何他们所面对的方向仍然如此。库克塔约尔人不仅已经知道方向（通常比我知道得更清楚），而且会自发地利用空间取向来给出他们的时间表示。

操不同语言的人的时间观念还有其他方式的不同。例如，说英语的人在谈论时间时喜欢运用水平空间比喻（例如，“我们面前的这个最好”“我们后面的那个最差”），而讲汉语的人则喜欢用垂直比喻（例如，“下个月”“上个月”等）。讲汉语的人要比讲英语的人更经常用纵向概念来谈论时间，那是不是说讲汉语的人考虑时间问题时也要比讲英语的人更经常运用纵向概念呢？想象有这么一个简单实验：我站在你旁边，指着你面前的一个空间点对你说：“这个点，看到了吧，指今天。你认为昨天应在什么位置？明天又会在什么位置？”如果说英语的受试者被要求这样做，他们几乎总是沿水平方向点。但说汉语的受试者则经常是垂直地点，其比例要比说英语的受试者高上7～8倍。[29]

即使是基本的时间知觉方面也会受到语言的影响。例如，说英语的人谈论时间时更愿意用长短来修饰（例如，“这是一次简短的谈话”“会议没有多久”），而说西班牙语和希腊语的人则更愿意用数量来修饰，即更多地采用“很大”“大”和“小”等字眼而不是“短”和“长”来形容时间间隔。对时间间隔进行估计的这种基本认知能力的研究显示，说不同语言的人在进行时间估计时所用的隐喻模式是不同的。（例如，当被问及估计一下时间时，说英语的人较易受到距离信息的困扰，他们往往用测试屏上较长的线来作为对较长时间的估计；而说希腊语的人则更容易受到数量的困扰，他们用屏幕上容器的充满程度来估计时间。）[30]

这方面的一个重要问题是：这些差异是由语言本身所造成的呢？还是由文化的其他方面造成的？当然，说英语的、说汉语的、说希腊语的、说西班牙语的和库克塔约尔人等的生活方式千差万别。我们如何知道是语言本身而不是各自文化的其他方面造成了思维的这些差异呢？

回答这个问题的一个办法是让人们学习一种新的说话方式，然后观察他们运用这种方式说话时思维特点是否有所改变。在我们的实验室中，我们教说英语的受试者用各种不同的方式来谈论时间。譬如在一次这种研究中，我们让说英语的受试者用大小的比喻（像希腊人那样）来

描述时间（例如，电影要比喷嚏大），或用垂直的比喻（像说汉语的人那样）来描述事件的顺序。一旦说英语的受试者掌握了这些新的谈论方式，他们的认知能力就开始表现得像个说希腊语的人或讲汉语的人。这表明，语言模式在塑造我们的思维方式时确实能够发挥作用。[31] 从实用上说，这意味着当你学习一门新的语言时，你不只是在学习一种新的说话方式，而且是不可避免地在学习一种新的思维方式。

除了空间和时间这样的抽象或复杂的思维域之外，语言也对视知觉的基本方面——例如区分颜色的能力——有影响。说不同语言的人划分色谱的方式是不同的：一些人就比另一些人区分得更精细，操不同语言的人区分的界限往往不一致。

为了测试色彩用语的不同是否会导致色彩认知上的差异，我们比较了讲俄语的和讲英语的人对蓝色色调的鉴别能力。俄语里没有一个用来涵盖英语里“蓝色”所指称的所有蓝色的词。俄语对浅蓝色（goluboy）和深蓝色（siniy）做了强制性的区分。这是否意味着在俄国人看来siniy色看起来与goluboy色很不相同呢？实验数据表明确实如此。讲俄语的人能用不同的俄语单词命名来更快地区分出两种蓝色（即一种称为深蓝，一种称为浅蓝）。而对说英语的人来说，所有这些色彩都归属于同一个词：“蓝色”，而且在实验规定的大脑反应时间内想不出可比较的差别。

但当受试者被要求执行口头干预任务（念一串数字）同时进行颜色判断时，说俄语的这种优势消失了。而当受试者被要求执行同样困难的空间干预任务（记住新的视觉模式）时，则这种语言优势还存在。执行口头任务时这种优势的消失表明，语言通常参与极为基本的知觉判断——正是语言本身造成了说俄语与说英语的人之间的这种差异。当讲俄语的人被口头干预任务阻断了这种正常语感时，说俄罗斯和说英语的这种差异就消失了。

即使是语言的那些被认为琐屑不起眼的方面也对我们如何看待世界有着深远的潜意识的影响。以语法上的词性为例。在西班牙语和其他罗曼语里，名词不是阳性就是阴性。在许多其他语言里，名词被分为更多的性（“性”在这里仅具分类的意义）。例如，澳大利亚的一些土著语言中有多达16种性，包括猎枪、犬齿、有光泽物体等类。正如著名的认知语言学家乔治·拉克夫（George Lakoff）说的那样，这些词性可归为“妇女、火和危险的事情”。

语言在语法上有词性意味着，属于不同性别的词需要进行不同的语法处理，属于同一性别的词则遵从相同的语法处理规则。根据名词的性别，语言可以要求说话的人改变代词、形容词、动词词尾、所有格和单复数等。例如，俄语里说："我的椅子旧了（moy stul bil'stariy）"，你需要将句子中所有词的性与"椅子"（stul）的性保持一致，而这个词在俄语里属阳性。因此，"我""是"和"旧"这三个词必须使用阳性形式，就像你谈到生物上的阳性譬如"我祖父很老"这个句子里用的词的阳性形式一样。但如果现在不是说椅子，而是说床（krovat），这个词在俄语里属阴性，这时说话时用词的形式就与谈到你祖母时一样，"我""是"和"旧"这三个词需要用阴性形式。

俄语语法里椅子是阳性而床是阴性，这是否会使讲俄语的人认为椅子更像男人而床更像女人呢？事实证明确实如此。在一项研究中，我们要求讲德语的和讲西班牙语的受试者描述这两种语言中不同性别的对象。按语法性别预测，他们给出的描述具有差异。例如，当被要求描述"钥匙"——这个词在德语里是阳性而在西班牙语里是阴性——一词时，说德语的更可能使用诸如"硬的""重的""参差不齐的""金属的""锯齿状的"和"有益的"等词，而说西班牙语的则更有可能用"金色的""复杂的""小的""可爱的""发光的"和"单薄的"等词来形容。当描述"桥"一词时，这个词在德语里是阴性而在西班牙语里是阳性，于是说德语的受试者多使用"美丽的""优雅的""脆弱的""和平的""漂亮的"和"修长的"等词来修饰，而说西班牙语的受试者则用"大的""危险的""长的""结实的""坚固的"和"入云霄的"等词来修饰。即使让他们用语法上没有阴性阳性之分的英语来测试，结果仍然是这样。这些结果的相同模式还出现在完全非语言的任务上（例如，评价图片之间的相似性）。我们还可以证明，正是语言本身塑造了人们的思维特点：讲英语的人学习了新的语法性别系统后，在心里描述对象时就会受到语法上性的影响，其影响的特点与讲德语和讲西班牙语的具有相同的方式。显然，即使是语法上一丁点儿小的偶然性，譬如将一种语法词性任意分配到一个名词上，都会对人们指称具体对象时的思维产生影响。[32]

实际上，你甚至不需要到该实验室就能看到语言的这些影响，你用自己的眼睛在艺术画廊就可以看到这一点。看看一些著名的人格化艺术的例子——其中抽象的实体，如死亡、罪恶、胜利或时间等，是如何通过人的形式被表现的。艺术家是如何决定用男人还是用女人来表现死亡——譬如说——或时间的呢？事实证明，这种人格化表现，无论是男性

还是女性形象的选择，有85%的把握可通过艺术家母语单词的语法性别预测出来。例如，德国画家更有可能用男人来表现死亡，而俄罗斯画家则更有可能用女人来表现死亡。

语法的这种怪异性（如语法有性别）能够影响我们的思维，这一事实具有深远影响。这种怪异性渗透在语言的各个方面，例如，性别就适用于所有名词，这意味着它将影响到由名词指称的任何对象。这是多么大的一群！

以上我描述了语言是如何影响到我们对空间、时间、颜色和物体等的思考。其他研究发现，语言还影响到人们如何解释事件、因果关系、掌握数字、了解材料性能、认知和情感体验、推测别人的思想、选择承担风险，甚至我们的择业和择偶方式。[33] 总而言之，这些结果表明，语言的过程遍及思维的最基本方面，无意识地塑造着我们的思维方式，从认知和感觉的基本单元到最抽象的概念和我们生活中的重大决策，无一能够摆脱。语言是我们作为人的经验的核心，我们讲的语言深刻地影响着我们的思维方式、我们观察世界的方式和我们的生活方式。

第11章
记忆增强，记忆擦除：我们过去的未来

◎萨姆·库克

萨姆·库克（Sam Cooke）

神经系统科学家，他的工作涉及两个问题：学习时我们的大脑是如何改变的？大脑如何保持这种变化并存储为终生难忘的记忆？他在英国谢菲尔德大学学习哲学和实验心理学时开始对这些问题感兴趣。2002年，他从伦敦大学学院获得博士学位，论文题目为“运动记忆的产生”。在此期间，他澄清了通过联想学习获得简单运动技能的某些基本神经机制。毕业后他去了伦敦北部的国家医学研究所，并将研究方向转向情景记忆的神经基础研究。所谓情景记忆是指人对所经历的事件的细节的记忆。他还开始研究长时程增强现象，一种关于大脑记忆是如何通过突触改变来存储的模型。目前他是麻省理工学院的博士后，在这里他继续探索记忆生物学。

在21世纪里，人类很可能在内在经验性质的认识方面取得深刻变化。一旦我们理解了我们的记忆是如何形成、储存并在大脑中记忆的，我们就可以操纵它们——塑造我们自己的故事。我们的过去——或至少是我们对过去的回忆——可能成为一个可选择的问题。

我们已经见识了号称是益智药的各种“聪明”药物的出现。这些药物提高了我们的学习速度和记忆容量，我们也许可以发展出一种选择记忆的技术，来保持我们希望记住的并摒弃那些我们不希望记住的。最后，在遥远的未来，我们可能不必通过学习或实际生活体验来建立记忆。这些技术引入社会后会带来什么样的潜在的伦理后果呢？

在我们所拥有的所有素质里，只有记忆最好地将我们定义为一个个有别于他人的个体。例如，同卵双胞胎具有相同的基因，他人很难一眼

就将他们区别开来，甚至对那些认识他们的人都很难，但他们的记忆是唯一的。有了记忆我们才能在世界上生存。记忆不仅提供了我们对年轻时玫瑰色的回忆，还让我们懂得了什么是正确的，什么是错误的；让我们学会了如何走路、说话、系鞋带；让我们有了可周游世界的地图。记忆还使我们能够分辨敌友，分辨家人和陌生人，区分事实和虚构。没有记忆，我们不仅失去了自我；从我们自身的角度来看，我们甚至都不存在。

益智药与记忆增强

记忆随着年龄变大而衰退是我们这个时代的特征。人类的平均年龄，特别是发达国家的平均年龄，因为预期寿命的提高和出生率的降低而以惊人的速度上升。据2003年2月美国疾病控制中心的题为“公众卫生及老龄化”的报告透露，在世界范围内，65岁以上的人口数量有可能在2000～2030年期间增加一倍以上，从4.20亿增加到9.73亿；而世界总人口则以相当慢的增长率增长；老年痴呆症折磨着很大一部分老年人；在美国，仅患阿尔茨海默病的老年人就占65岁以上老年人的10%，而且年纪每大5岁患此病的风险增加一倍。身份丧失及其对病人家庭的影响不只是花钱的问题：这种精神衰退给全社会带来的经济负担是巨大的，并且将随着老年人口的增长变得更大。目前许多神经科学研究的目标都指向与老龄相关的记忆障碍的认知治疗。

用于治疗老年痴呆症的药品分为两类：一类是直接针对记忆障碍造成的记忆缺失，另一类则通过增强其他不相关的记忆过程来掩盖这种缺失。这第二类药物称为益智药（nootropics，这个词得名于古希腊词nous（精神）和tropein（转变）的合成），它也可以用来增强没患老年痴呆症的人的记忆。虽然记忆障碍的直接治疗非常重要，但益智药的研究也是制药公司特别感兴趣的一个领域，因为它的市场潜力更大。作为药品销售的第一种这类药叫吡拉西坦（Piracetam），商品名称叫益智丸（Nootropil）。尽管对其疗效和作用机制缺乏强有力的论证，但它非常受欢迎。潜在用户范围从期末考试的学生到试图在工作中出人头地的专业人士，不一而足。所有人都希望提高他们的记忆力和认知能力。

益智药本身可分为两类：①直接作用于信息存储和学习过程的药；②那些增强或模仿大脑的其他重要功能（如注意力、睡眠和奖励）从而

促进学习和记忆的药。第二类药中好些早已为我们所熟悉（尼古丁、咖啡因、葡萄糖、安非他明、可卡因），天然植物和真菌的提取物如银杏叶、人参和海得金（Hydergine，麦角生物碱甲磺酸盐类）也已被明确作为益智药销售。有些药物属于无心插柳，是在广泛用于治疗其他精神疾病后被确认为具有益智作用的。我们可举出莫达非尼和利他林作为两个例子，前者原本是用于发作性睡病的治疗，后者用于治疗注意缺陷多动障碍症（ADHD）。目前这两种药物都被那些没有上述障碍但想提高认知水平的人当作补品使用。让制药公司真正得到实惠的很可能是其他种类的益智药——那种直接作用于学习的分子过程的物质。人们希望，这种东西最好是比那些减少睡眠或增加注意力的药物的有害副作用更少。

我们距离完全解开记忆形成的秘密还有一段路程要走。大脑高速处理信息的能力以及能够瞬时变化以纳入新信息的能力使其成为大自然的伟大奇迹之一。大脑的处理速度是通过使用电信号（发送沿神经传导的冲动）和叫作神经递质的化学物质相结合来实现的，神经递质在称为突触的特化沟通点将信息从一个神经元传递到下一个神经元。具体来说，在突触处，神经递质漂移过称为突触间隙的小缝隙，并在另一边被叫作受体的专门收集站所收集，收集站再将化学能重新转换成电脉冲。这种突触传递过程是科学研究的一个主要焦点。

改变神经元之间沟通的方法有几种。在电传递过程中，发生改变的可以是释放的神经递质分子的数量，也可以是收集这些神经递质的受体分子的数量或效能。此外，突触形状或数量的变化将会对突触传递造成长期影响。这种修改称为突触可塑性。神经学家认为，记忆可能正是通过这种突触可塑性（在脑中形成沟通渠道）形成的。医药公司也对这些过程感兴趣，因为它们代表着潜在的益智药目标。尽管人们已做出很多努力，但迄今为止，直接用于提高突触可塑性的药物尚未面市。不过实现这一目标只是个时间问题。这种药物如果被证明是安全有效的，无疑将获得巨大的全球性销售市场。

选择性失忆

创伤后应激障碍（PTSD）是这样一种记忆疾病，它属于阿尔茨海默病系的另一极端：患有这种疾病的病人的大脑的信息存储是完好的，

从而使特定的创伤事件深深地烙在病人的心中，使他们因恐惧而变得极度虚弱。这种病证通常出现在被强奸的受害者和战争归来的士兵身上，是当前医学研究的另一个重要方面。对PTSD的一种有效的治疗方法是在心理上擦除恐惧的原因，同时保留对人类经验的美好回忆。但如果人类的记忆依赖于相同的生物学机制，我们如何才能有选择地进行记忆而不会影响到其他方面呢？

人们早就知道存在所谓的遗忘剂——一种在学习后的短时间内阻止形成长期记忆的物质。我们大多数人可能都有过这样的经历，晚上喝了太多的酒后记忆会丧失。吸食大麻的人也都知道这种毒品对他们的记忆有负面影响。但这种治标不治本的方法不适合PTSD的治愈，因为这种疾病的病因是突发事件，我们不可能事先获知而将自己灌醉。因此这里需要的是一种可以在事后实施治疗的药物，它可以影响到记忆的巩固过程。记忆的巩固通常发生在学习后的一段有限时间内，在此期间记忆会从早期脆弱的状态转变为后期永久性的状态。我们的神经生理在记忆巩固期内会出现某些重要的变化，使得记忆所基于的神经活动改变稳定。正是这些重要变化使我们能够保留终生难忘的记忆。

新蛋白质的合成似乎是其中的一个重要因素。人们差不多在50年前就已经知道，抗生素不仅能够阻止细菌产生新蛋白质来灭菌，而且对各种动物（从金鱼到人）采用不同的剂量可以使动物产生遗忘。[34] 这些结果第一次以科学证据表明，我们的记忆是由新蛋白质建立起来的。人们认为，不是暂时性记忆起因于突触上现有蛋白质的改变，而是长期记忆则需要这些蛋白质更新。但使用蛋白质合成抑制剂来治疗PTSD不太可能成为一种可行的方法，因为细胞的几乎所有主要功能，从分裂到修复，全都要依靠蛋白质的合成，因此采用抑制剂会带来损伤性副作用。

另一种做法是将目标锁定在大脑中调解情绪应激反应的系统。相关物质包括肾上腺素以及与此密切相关的去甲肾上腺素（一种在紧张情绪体验过后大量释放出的主要神经递质）。实验表明，在学习后记忆巩固的关键期，使用β-阻断剂（一种临床上用于治疗心脏疾病和怯场的药剂）来阻止去甲肾上腺素与受体的结合可以阻止大鼠和人的长期情绪记忆。[35] 普萘洛尔——β-阻断剂中的有效成分——目前已被大量应用于治疗PTSD的临床试验。

但这种疗法有其局限性，因为普萘洛尔必须是在突发事件过后的一

小时左右施治才有效。而士兵在战斗情况下，你可以设想，士兵带上这些药片，一旦出现引起PTSD的事件后将立即吃药。在其他更意想不到的情况下，例如强奸或车祸，这样的治疗可能更不容易获得。但最近对情绪记忆的研究暗示了解决这个问题的一种办法。某些记忆形式并不简单遵循从脆弱到稳定的单向时序。事实上，由于存在所谓重新巩固（reconsolidation）的现象，某些情绪记忆，无论什么时候被唤起，都会重新经历一段时间，在这段时间里，采用普萘洛尔或蛋白质合成抑制剂很容易抹去这段记忆。[36] 因此，擦除记忆可以在任何时候进行，只要在服用普萘洛尔之前让病人回忆所发生的事件以便进入重新巩固状态即可。这种疗法还没有完全在人身上试用，要采用这种技术来治疗PTSD可能还要等上一段时间。但不论怎样，它最终可能会广泛应用于这类疾病以及其他一些记忆障碍疾病，如恐惧症或强迫症。

这种“回忆”疗法很可能让那些没患有这种疾病的人感兴趣。正如使用益智药使正常人的记忆增强一样，记忆擦除技术将带来深刻的社会影响——一种在电影《美丽心灵的永恒阳光》（*Eternal Sunshine of the Spotless Mind*）里表现出来的场景：那种不需要的失败的爱情记忆被抹去，人们可以对世界保持一种理想化的观点。难道在未来我们真的能够将旧情人的一切从我们的头脑中完全抹去吗？

无需学习即获得记忆——玛丽莲·梦露实验

横在记忆科学面前最具挑战性的障碍是创建一个对从未发生过事件的记忆。这被形容为玛丽莲·梦露实验——之所以如此命名，是因为许多年长的男性科学家可能愿意为自己创造一个与20世纪最富性感的女神共度良宵的美好记忆，如果有机会的话。尽管这个实验的名称显得轻浮，但如果它能够获得成功，那将是一个巨大的成就——它最清楚地给出了我们关于记忆的生物学基础理解的证据。

关于大脑如何存储信息已有各种强有力的假设。当然到目前为止，这些仍然是假设。我们知道，突触可塑性所涉及的分子很多都与记忆有关。我们还知道，突触传递效能的改变往往伴随着新记忆的形成。但是这些研究结果只是间接证据，它们揭示了现象之间的相关性而不是因果关系。为了有力地证明突触可塑性的形成和记忆存储之间存在着的因果关系，科学家们必须证明，突触可塑性对于学习的发生和记忆存储是必

需的。要做到这一点，科学家必须在不损害我们其他方面生理功能的情形下有选择地禁用突触可塑性，然后他们必须观察到记忆的丧失。在某些方面，上述记忆擦除实验可以做到这一点，但是，即使我们能够证明其必要性，但仍存在一个突触可塑性是否充分的问题，就是说，它必须是与学习和储存记忆有关的唯一过程。为了检验这一假设，我们必须证明，化学突触功效的变化足以创建记忆。这就是玛丽莲·梦露实验要取得的成果。

必要性和充分性两方面的证明界定了科学证据的过程。不幸的是，检验这两个判据的实验也最难进行。就突触可塑性和记忆这一课题而言，科学家还没能做到这一点，这样的实验还在进行中，因此突触可塑性理论仍然是一种假说。做这种实验的困难有两个：第一，科学家们必须在人脑的10^{15} 个突触中准确确定与特定记忆的形成和存储有关的子集——既不多也不少；第二，他们必须能够以一种模拟自然生理的方式来人为地改变突触。目前，这类实验正在动物（大鼠和小鼠）身上进行。与人相比，这些鼠类的突触数量大大减少，并且我们可以采用植入的方法来操控突触的强度。

对啮齿动物的记忆进行研究采用的是训练它们走精心设计的迷宫。动物利用水池周围的空间提示物来学习寻找不透明水面下隐藏的平台，或通过特定地点或声音与随即发生的电击之间的关联来使它们形成恐惧记忆。其他一些迷宫测试和训练仪器则用来检测啮齿动物的各种记忆，每一种记忆依赖于不同的大脑回路，科学家将这每一种记忆看作人类特定记忆类型的模型。我们开始了解脑的哪一部分区域传递着哪种类型的记忆——情景记忆（对事件发生的时间和地点的记忆）、情绪记忆（对可怕事件或褒奖事件的记忆）、程序记忆（对获得的技能的记忆）等。光学成像和显微镜等新技术让我们能够实时观察突触在学习过程中所发生的变化。一旦我们能够观察到大脑在学习过程中的分子相互作用，并在足够精细的水平上测量神经活动的有关模式，我们就可以最终检测突触可塑性在记忆过程中的因果作用。

在不久的将来，玛丽莲·梦露实验的最佳逼近实验将能够在大鼠学习将地点与电击关联起来的过程中实时观察到大鼠大脑中发生的确切的模式变化。然后我们可以尝试着抹去这段恐惧记忆，这种抹去的效果可以通过观察动物在原先受到电击的地点呆着不动所花时间的长短来衡量。通过使用普萘洛尔将这段记忆完全擦除，应能使大鼠在这个地方完

全不以不动状态停留，而是自由往来，就像是第一次来到这里。然后才谈得上重新安装记忆，即人为地复现我们在动物第一次受到电击时观察到的确切变化。如果动物确实又表现出呆立不动，这将表明大鼠的临震恐惧记忆已成功得到重新安装。

这一系列实验证据显然不能完全等同于真正的玛丽莲·梦露实验，因为动物亲身经历了最初的主要事件。尽管如此，它们在概念上是类似的，并有可能实现神经科学家要证明的特定生物过程的充分性的目标——眼下这个目标就是突触可塑性对于记忆形成的充分性。人们希望，通过在不同动物身上就特定内容的记忆进行实验观察，对其中的共性形成一种更广泛的理解，并最终使我们能够进行真正的模仿实验，即让记忆在一只完全天真的老鼠脑中复现。

设计记忆和科学发现的滥用

我们是否真的能将这种模仿实验移植到人身上，并尝试着进行玛丽莲·梦露实验完全是另外一个问题：首先，我们将不得不问自己进行这种努力是否值得。

历史表明，社会为达到任何可能的目的最终会采用一切可获得的知识和技术，不管这种技术手段是多么令人讨厌。各种科学发现最初都推进了我们对宇宙的理解，但随后很快就被付诸令人不齿的用途。从核裂变的发现到原子弹的出现是这样，从弗朗西斯·加仑的优生学到约瑟夫·门格尔在第二次世界大战中实施的人体实验也是这样。这种从发现到应用的进展在提高生活质量方面表现得也很明显，整容手术、伟哥和类固醇的泛滥就是例证。整形外科最初是出于救治严重烧伤病人这一仁慈目的发展起来的，伟哥则是作为一种治疗心绞痛的药物开发的，合成代谢类固醇的最初用途是重建伤后的肌肉以及治疗生长性疾病。但这三种疗法无一例外地都被用作他途来获利。人们或许会辩解道，这些生活方式的改进并不都应受到谴责，但这里想强调的是社会有必要探讨新的技术方法的所有潜在用途。

最先作为治疗阿尔茨海默病和其他形式老年痴呆症的益智药能找到更广泛的用途几乎是无可置疑的，譬如提高青少年和正常人的记忆。记忆擦除技术也可能成为普通民众的现实。人工记忆的形成目前仍是幻想，但总有一天会实现。随后设计记忆的丑陋前景就该出现了。通过长

期服用益智药，我们可能会变得更加聪明。通过选择性记忆重建技术，我们还可以创建一个干净的内心生活。难道我们一定要形成一种意识经过美化，负面经验被剥夺，到处充斥着正面的虚假记忆的局面吗？这些技术真的能成为强有力的额外的精神控制武器，能够消灭我们记忆中的政治恶行并植入经过编辑的清洁记忆吗？

这种反面乌托邦显示了科学家的一种道德困境：当我们知道下述探索会从根本上改变多年来形成的社会结构时，我们还该不该为了治疗像阿尔茨海默病这样的认知失调症而去研制益智药？该不该尝试选择性地抹去由PTSD引起的恐怖记忆？该不该探索玛丽莲·梦露实验以最终检验我们关于记忆的生物学基础的假设？有人会说，对研究可能产生的影响作出道德判断，这不是科学家的事情，而是政府或是选民的责任，该是他们来决定哪些科学研究值得公共资金投入，哪些不。还有一种观点认为，任何新知识或技术引入公共领域所带来的复杂影响是很难预测的，我们不能为了回避未来的负面作用而抛开眼前利益。只有时间能告诉我们益智药和记忆重建技术是否对社会有益。但过去的经验表明，正像其他科学发现一样，我们很快就会使用和滥用这些成果，因为不管我们喜欢与否，我们已经登上迈向享受益智药、永恒的阳光和玛丽莲·梦露的“智能”新世界的征程。

第12章
想象的极端重要性

◎蒂娜·斯科尼克·韦斯伯格

蒂娜·斯科尼克·韦斯伯格（Deena Skolnick Weisberg）

耶鲁大学心理学系的博士生，攻读方向是发展心理学。她于2003年获得斯坦福大学认知科学学士学位，并在耶鲁大学先后获得了科学硕士和哲学硕士学位。韦斯伯格的研究主要集中在对非真实场景——尤其是故事、假想游戏和非真实情形——的创造和表达的认知技能，以及在儿童成熟过程中这些技能是如何发展的等方面。她还研究儿童如何学习新的形容词，研究成人误解科学发现（特别是神经科学方面的发现）的各种情形。她的研究已发表在多种学术期刊上，包括认知神经科学专业杂志《认知》和《科学》。

关于我们的认知能力的一个不寻常的事实是：我们不只是生活在真实中。我们可以回到过去（通过记忆回到我们自己的过去，或通过历史回到世界的过去），或展望未来，或神游幻境，并想象出那些我们明知其不真实也永远不会实现的各种可能性。

即使是非常年幼的儿童，也都能作出这种超越现实的幻想。这种能力的最早期的表现就是“假扮”游戏。孩子可以在没有任何茶点的情形下开茶话会，或一边拿着铅笔在桌子上游动一边学着汽车叫，他完全知道铅笔不是汽车。事实上，研究人员发现，两岁的孩子，作为这种假扮游戏的新手，已经懂得将假想的水泼在玩具熊上来浇湿熊宝宝，即使实际上没有水可浇，玩具熊也不会湿。[37]

随着孩子年龄的增长，他们对现实与幻想之间差异的理解力逐步深化。他们开始用一种成熟的方式来进行假扮游戏，感兴趣的故事也不再是他们自己的创造。最迟长到4岁，往往还要更早些，孩子们已经熟练

掌握了真实和幻想之间的区别。他们报告说，像巫师和小仙女这样的生物并不是真实存在的。他们知道小熊在厨房做饭的图片描写的并非真实的事情，虽然小熊和厨房都是真实的。[38] 他们还懂得现实与幻想间差别的本质，因为他们知道真实的甜饼是可以吃的，而虚构的甜饼是不可以吃的。[39]

即使孩子们知道只有两类对象：一类是真实的，一类是虚构的，他们仍然明白所有这些区别。如果情形确是这样，那说明他们已能够正确地理解所有真实的物和人都属于现实世界，并相信所有虚构的物和人都属于同样虚构的世界。

下面这一点往往不是成年人所愿意相信的。成人通过不同的故事编造出各种虚构的世界，通常没指望不同故事里的角色互动。关公战秦琼那样的交叉故事——例如《怪物史莱克》或《新蝙蝠侠》——之所以显得生动有趣，正是因为它们出人意表，它们不遵从那种隐含的规则，即不同的故事是相互分离的。在正常情况下，故事中的人物都待在自身适当的虚构世界里。

我设计了一组研究实验用以了解儿童是否像成人那样能够理解幻想王国里的这些差别，还是他们只能理解真实与幻想之间的区别。我向4～5岁的儿童问了3个问题，都是关于他们熟悉的角色的，像蝙蝠侠和棉球方块等。首先，我问他们是否明白这些人物其实是虚构的。（“你认为蝙蝠侠是真实的还是捏造的？”）其次，我问他们是否明白在同一个故事里，虚构人物之间彼此认为对方是真实的。（“蝙蝠侠对罗宾是怎么看的？蝙蝠侠认为罗宾是真实的还是虚构的？”）最后，为了测试儿童是否能够将故事组织成不同的世界，我问他们不同故事里的人物可不可以彼此接触。（“蝙蝠侠对棉球是怎么看的？蝙蝠侠认为棉球是真实的还是虚幻的？”）

儿童对所有三组问题的反应与成年人对同样问题的答复没有两样。像成年人一样，孩子们认为，蝙蝠侠是虚构的，但是蝙蝠侠和罗宾彼此之间认为对方是真实的，而蝙蝠侠和棉球都认为对方是虚构的。[40] 关于这些反应的一个显著特点是，儿童（成人也如此）可能不会明确想到各人物之间的关系，但当我让孩子们归类这种关系时，他们做起来一点都不困难。儿童和成人一样为他们知道的故事设计了组织结构，即使他们可能没有意识到他们为什么要这样做。

这一结果促使我想知道孩子们在假扮游戏中是否也同样会区分清楚。假扮游戏与故事一样，都是表现非现实的情形，都涉及人物和一系列行为——但与故事不同的是，在假扮游戏中，儿童有更多的角色体验，而且可以直接控制角色。那么到底是什么因素决定了儿童能够区分假扮游戏世界中的相同或不同点呢？

为了回答这个问题，我和我的研究助理创建了两个假扮游戏，所设的场景类似于先前研究中提出的两个故事情形。我们用一组彩色积木作为两个游戏中的假想物。例如，在第一个游戏中，玩具熊需要洗澡。孩子可以用积木作为假想的肥皂来给小熊洗澡。在第二个游戏中，布娃娃需要午睡，孩子可以拿第二块积木给布娃娃当枕头。

现在问题的关键是，儿童是将两个游戏看作是不同的世界呢，还是将两者看作是可以互通的？在这里，我想测试孩子是否会将一个游戏里的东西移到另一个游戏里。因此我对孩子说，现在小熊洗过澡要睡觉了，小熊需要一个枕头，你给小熊拿个枕头。这里孩子的反应可能是两种：从娃娃游戏中取过枕头或拿一个新的积木当小熊的枕头。前者的选择可能表明，孩子对两个假扮游戏不做区分——就是说，一个游戏里的物体可以毫不犹豫地移到另一个游戏里。后者的选择将意味着孩子对两个游戏做出了区分，他不愿将相关物体从一个游戏移到另一个游戏里。研究表明，几乎绝大多数孩子都会选择一个新积木块做枕头而不是将另一个游戏里的枕头拿走。而且两个游戏不论是先后进行还是同时进行，他们的反应都一样。

因此似乎是，孩子对待假扮游戏与他们听故事时是一样的，两种类型的虚构情形具有相同的组织结构。存在共同结构这一点表明，人在故事和假扮游戏里采用的思考方式是同一种认知机制，我称之为假设—推定（what-if）机制。顾名思义，这一机制是一种检测认知的手段，我们可以通过设问“如果……会怎样？”来探索现实中不存在的各种可能性。通过发挥想象、设立场景并在此背景下进行演绎就可以实现这种检测。有时我们还可以借助于文学作品和电影来实现这一点，我们可以按照作品的思路继续演绎下去来探索各种可能性。如果我一转身就隐形了那会怎样？如果真有一颗魔戒会怎样？如果这颗魔戒落入霍比特人[41]之手又会怎样？如果那个名叫郝思嘉的年轻高傲的姑娘生活在南北战争时期的美国南部会怎样？在所有这些情形下，为了将我们自身从当前现实中抽离出来，我们需要做的就是问“如果……会怎样”。

我认为我们对所有涉及现实以外情形的想象都有赖于这种机制的支持，无论是听故事、玩假扮游戏、做白日梦、想象未来或试图找出过去发生的什么事情，均不例外。在所有这些情形下，我们需要建立一个超脱现实的代表物。然后，我们知道我们正游离于现实之外，我们也知道这些代表物不是真的。在现实中，这是一支铅笔；但在假扮游戏里，它就是一辆汽车。在现实中，不存在布鲁斯·韦恩这个人；但在故事里，布鲁斯·韦恩不仅存在，而且过着一种打击犯罪的超级英雄的秘密生活。当我们设想另一种过去（反事实推理）或规划着未来（假设推理）时，采用的是完全一样的推理过程。在现实中，事态只以一种方式发展；但在我们的想象里，事件则可能完全按另一种方式发展。在现实中，我们只能看到眼前的一切；但在想象中，我们可以提出各种可能的未来图景，以帮助我们预测将会发生什么，最佳的行动方案是什么。我把所有这些类型的陈述称为虚构的世界，用以突出其两大属性：它们不是真实的，而且我们知道它们不是真实的。

我的研究的主要目标是找出假设推定机制的性质以及它是如何使我们能够创造并理解虚构世界的。在这两项研究中，至少有一个印记已显露出来：多重虚构世界是彼此分开的。这些研究在故事和假扮游戏里都发现存在这种多重世界结构，我希望在今后的研究中，我们从反事实推理和假设推理中也能找到这种结构。

除了多重虚构世界彼此分开这一点外，假设—推定机制的另一项工作是探寻现实与虚构的边界。从某种意义上说，这条边界是刚性的。像蝙蝠侠这样的人物是虚构的，不可能转移到现实中。假扮游戏里当作甜饼的积木也不会被误认为是现实中真正的甜饼。但在其他方面，这条边界要模棱两可得多。譬如，我们是带着现实世界的许多特征进入虚构世界的。蝙蝠侠可以成为打击犯罪的超级英雄，时间旅行也是可能的，但二加二仍然等于四，水在冷冻时仍然变成冰，一件东西是蓝色的它就不可能同时又不是蓝色的。各种数学、科学和逻辑上的规律仍然有效，不论这种虚构的世界有多大的不同。因此，假设—推定机制必须能够将现实世界中的事实和规则以适当方式投射到虚构的世界。

此外，当我们从虚构世界里出来回到现实中来时，我们带回来虚构世界里的许多特征，这么说听起来似乎有些奇怪，但考虑到我们从读书、看电影和电视节目中学会了很多关于现实世界的许多东西（如果没有这些媒体，有些重要的知识我们可能永远学不到），就一点也不奇怪

了。当我们阅读或观看《奥赛罗》时，我们懂得了嫉妒具有破坏力，即使我们从来没有被嫉妒逼得发疯，而且可能永远也不会有。当我们观看电视剧《犯罪现场调查》时，我们学会了怎样收集和分析DNA证据，即使我们在现实生活中从来没有见过或使用过这些技术。如果有人告诉你，他打算通过阅读托尔斯泰的小说来寻找19世纪俄国的生活特征，你不会感到惊奇。从小说里学习如何应对现实问题这种事情无时无刻不在发生，无论是了解情感的真实性，还是感受一个国家或一个时代，或拾取某些实用信息，无不如此。

正是虚构世界的认知特点使得它对我们非常有用，对成长中的孩子就更有用了。有许多现实情形是我们无法直接探索的，对儿童更是如此。但通过我们的想象，我们无需去实际经历仍然可以学到很多东西。事实上，一些研究人员推测，这正是假扮游戏的意义所在：让孩子们有更多机会扮演其他社会角色，或在一个安全、独立的环境下探索他们个性的其他方面。儿童从假扮游戏里可能不像听故事那样能够了解到事情的实际情形，但他们可能会学习社交经验和情绪体验，譬如在某些情况下如何互相交往，或做一件现实生活中从未经历过的事情是什么感觉，等等。

每天都有需要我们去解决的问题。我们需要借助于假设—推定机制来处理。譬如汽车发动不起来了，我们必须找出原因。我们可以设想一种虚构的情形：是电池没电了吗？是变速器坏了？对于我们想象得出的每一种可能性，我们用假设—推定机制来构造一个情景，如果这种可能性是真实的，那么接下来会发生什么。然后我们回到现实中来，利用想象场景得到的信息来检验这种可能性。如果是电池没电了，我可以采用助推启动方式来发动汽车；如果是变速器出了毛病，那么助推启动就不会起效。甚至一些简单的计算也需要发挥想象力，需要采用假设—推定机制，需要跳出现实外。现实不提供任何帮助，汽车不会自己启动，而这一切信息我都已掌握。为了让汽车启动，我可以独自一人竭尽各种可能情形以便从现实中获得更多的信息，但是这不是一种实际的收集数据的方法，而是通过暂时搁置事实，发挥想象，并从中找出错误。我可以最佳地利用我的时间和资源来解决我所面临的问题。

假设—推定机制的这种功能对于儿童更是必不可少，因为现实中会有那么多的问题让他们难于理解。他们不可能为了掌握事情的要点而去尝试每一种可能性。像成人一样，他们需要跳出现实，通过想象来工

作：如果做了某件事，结果会怎样。如果结果有利，他们就会继续尝试；如果结果不对，他们就会设想另一种可能性并采取不同的行动，他们不会把时间浪费在行不通的事情上。

假设—推定机制基本上是一种学习引擎。虽然它的主要功能是让我们超脱现实，但这些虚构的场景为我们提供了认识现实的新的窗口。玩游戏或听故事不只是很有趣，更重要的是在此基础上，假设—推定机制发挥着从这些虚构世界里获取对现实世界有用的信息的功能，这种机制扩展了我们的实际知识、情感体验和社会知识。关键是要确保它从中提取的是正确的信息。这是我以后要研究的主要问题：我们如何知道从虚构世界带回到现实里来的信息是适当的？我们自动地就知道这一点吗，正如我们自动地就知道如何创建虚构世界那样？抑或是需要经过经验和学习呢？

我相信答案在于两者之间。儿童可能内在地就知道他们可以从虚构场景中获得在现实世界里有用的信息。正是这一点使他们能够在他们所设计的未来世界里采取行动，也正是这一点让他们在假扮游戏世界里学着处理现实世界的问题。但成年人从虚构世界取得的东西与孩子的所获还是有细微差别的。举例来说，哈利波特不是真实的，我们不会将这个角色带到现实里来。这一点连孩子都知道。但是，他在英格兰所居住的房子——Privet Drive 4号——到底有没有呢？这条街在现实中存在吗？他的导师邓布利多（Dumbledore）教授肯定是虚构的，但现实中有没有类似的人呢？也许会有，但这种类似表现在哪方面呢？真实的人不会有邓布利多那般神奇能力，但他的其他身体或心理特征在现实中有潜在的原型吗？

成年人对这些问题的微妙的直觉不是孩子能够分享的。儿童的直觉往往比成人更偏激，也就是说，他们可能会把太多或者是太少的东西从虚构带到现实中来。儿童对虚构与现实之间界限的认识可能比成人更模棱两可，他们可能会将虚构世界里的几乎所有东西搬到现实中来，因此相信那些在现实世界里不存在的古怪的东西和规则会起效。相反，他们可能比成人更看重虚构与现实的分离。他们不明白，即使是在故事里，许多关于人们之间的相互交往和物理现象其实在现实中是很容易发生的。

我希望能够阐明这些问题。但现在有一点是清楚的：那就是假设—

推定机制对于儿童十分重要。这种基本的想象力技能，即让我们能够暂时跳出现实去发挥各种可能性的能力，是儿童了解世界的主要工具之一。

第13章 脑时间

◎戴维·伊戈尔曼

戴维·伊戈尔曼（David M.Eagleman）

在莱斯大学和牛津大学获得英美文学学士学位，1998年在贝勒医学院获神经科学博士学位。他目前担任贝勒医学院认知与行为实验室主任。这个实验室的长期目标是时间知觉的神经机制研究。伊戈尔曼还领导着贝勒医学院关于法律、脑和行为的创新计划，其目的是探求神经科学领域的新发现是如何改变我们的法律和刑事司法系统的。伊戈尔曼还是下述两本书的作者：《萨姆：晚年的40件轶事》（Sum：Forty Tales from the Afterlives ），以及《可塑性：大脑如何重新配置自身》（Plasticity：How the Brain Reconfgures Itself ）。

从某种程度上说，蒙古军队指挥官忽必烈（1215—1294年）意识到他的帝国变得如此广阔，以至于他永远无法看清楚这个帝国到底覆盖了多大区域。为了解决这个问题，他委派使者前往帝国的遥远边疆去传回他所拥有国土的最新信息。由于他的这些使者需要从不同的距离带回信息，传递的速度也不一样（取决于天气、战争状态和使者自身的健康状况），因此这些信息是在不同时间抵达的。虽然还没有历史学家考虑过这个问题，但我想象大汗一定不得不反复面临这样一个问题：帝国发生的这些事件的顺序到底是怎样的？

毕竟，你的大脑被静静地包裹在颅腔之内。它与外界的唯一联系是通过沿神经束高速路传递的电信号。由于不同类型的感觉信息（听觉、视觉、触觉等）是由不同的神经结构按不同速度来处理的，因此你的大脑面临着一个巨大的挑战：复现外部世界的最佳方式是什么？

将时间比作河流——均匀流动，永远向前——的时代已经过去了。

如同视觉一样，我们对时间的知觉是由大脑构建的，这在实验上验证起来相当容易。我们都知道有视错觉，即事情看上去与其实际情形不同，但很少有人知道还存在一种时间错觉。一旦你开始寻找时间错觉，你会发现它们似乎无处不在。在电影院里，你将一系列静态图像感知为一幅流畅的场景。或许你已经注意到，每当一瞥时钟秒针的一刹那，常常会感觉到秒针移动到下一个位置似乎需要比通常更长的时间，就好像时钟被暂时冻结了一样。

在实验室里你可以了解到更奇妙的时间错觉。在我们的眼球快速运动过程中——譬如在注视一个闪烁的亮光之后，或在一系列重复出现的图像中突然看到一个“怪物”时——我们对时间间隔的感知就会被扭曲。如果我们采取某种措施使你的运动性动作与你对这些动作的感觉反馈之间产生一些延迟，那么我们就可以使你感觉出你的动作和感受似乎出现了时间上的倒错。人的同时性判断在反复的非同时性刺激下会出现偏差。在自然界这个实验室里，像可卡因和大麻这样的毒品，或像帕金森症、阿尔茨海默病和精神分裂症这样的疾病，都会造成时间感觉上的扭曲。

试着做这样的练习：放下本书去照镜子。现在让你的眼睛反复来回运动：盯着镜子里你的左眼看，然后再盯着右眼看，然后再换到左眼，如此反复。按理说，当你将眼睛从一个位置移到其他位置时，是需要时间来移动并将视线落在某个位置上的。但这里怪了：你从镜子里永远也看不到你眼睛的移动。在你眼睛移动时到底发生了什么呢？为什么当你改变视线时你感觉不到时间的间断呢？（请记住，我们很容易看到别人的眼睛移动，因此不存在因眼球运动速度太快而看不到这种可能。）

所有这些错觉和扭曲都是你的大脑在构建时间表示时引出的后果。当我们仔细研究这个问题后，我们发现，“时间”不像我们想当然地认为的那样是一种单一的现象。这一点通过一些简单的实验就可以说明。例如：在不断播放的一系列图片中突然插入一张搞怪的照片，那么你会觉得这张搞怪照片的放映速度似乎较慢，虽然它的实际放映速度与其他照片并无二致。在神经科学文献中，这种效应最初被称为主观的“时间扩张”，并且这样一种描述引出了一个重要的时间表示问题：当感觉到时间扩张或收缩的那一刻，是不是一般意义上的时间感觉也放慢或变快了呢？譬如说，在你突然看到“怪物”时，是不是这时朋友说话的声调似乎也较低沉，就像播放速度突然变慢的唱机发出的声音那样？

如果我们的感觉的工作模式如同一部电影放映机，那么当某个场景的播放速度放缓时，其中的一切就都会变慢。譬如影片以慢速放映一辆警车冲过斜坡，这时不仅警车在空中会停留较长的时间，而且它的警报声也将变得较低沉，其灯光闪烁的频率也变得较低。另一种假说则认为，不同时间判断是由不同的神经机制产生的——它们经常同步，但这并不是必需的。因此警车可以在空中停留较长时间，但其警报声和灯光闪烁的频率不变。

现有数据支持第二种假说。[42] 时间错觉并不等同于电影放映时那样的统一的时间变慢。如同视错觉一样，我们对时间的感知依赖于各种单独的神经机制的协同，这些神经作用机制通常是协调的，但在特定情形下也可以是分开的。

这是我们在实验室中发现的情形。那么现实生活中的突发事件，如在新闻报道的车祸和抢劫事件中，人们感觉时间过得“很慢”是否可能有所不同呢？我和我的研究生钱斯·斯特森（Chess Stetson）决定将这个陈述变成一个真正的科学问题。我们的推理是，如果人在恐惧时感觉到时间是以一种统一的步调变慢，那么这个慢动作应该具有更高的时间分辨率——就像我们观看蜂鸟的慢镜头，我们可以观察到正常播放速度下看不清的蜂鸟的更精细的短暂行为，因为高速相机抓拍的速度要比蜂鸟拍打翅膀的速度更快。

我们设计了一个实验，在其中受试者只有在经历高时间分辨的情形下才可以看清一个特定的画面。我们利用的是这样一个事实：大脑视觉系统是在小的时间窗口内整合视觉刺激的，如果在一个时间窗内（通常为100毫秒）有两个或更多个图像被接收到，那么这些图像将被感知为一个单一的图像。例如，一种称为幻影转盘的玩具在碟片的一边可能是一幅鸟的图片，碟片的另一边是树枝的图片，当转盘高速旋转起来后，光盘两边就会快速交替，鸟看上去就像停在了树枝上。我们决定利用这种图片及其负片之间的快速交替作为刺激。受试者在图像交替速度很慢时辨认图像没有任何问题，但当交替速度加快后，就像鸟登上树枝一样，其结果是他们无法区分正片和负片。

为实现这一点，我们设计了一种仪器（知觉计时器），它可以按可调的速率交替地展示一个随机数字及其负数。我们用这台仪器测量了受试者在一般的放松情形下准确读出数字的阈值频率。然后，我们将受试

者吊到一个离地面15层楼那么高的平台上，知觉计时器则像手表一样绑在参与者的手臂上，只是交替显示随机数及其负数的速率要稍快于阈值频率。然后放手让受试者经历三秒钟的自由落体降落（绝对安全！）。在下落过程中，受试者读取数字。如果在自由降落过程中受试者经历了较高的时间分辨，那么他感知到的图片交替出现的速度就会变慢，他就能够准确读出通常状态下无法辨认的数字。[43]

结果如何呢？受试者在自由落体情形下读取数字的能力并不比实验室情形下更好。这并不是因为他们闭上了眼睛，或者不够专注（这一点我们有监测），而是因为他们终究无法像时间变慢情形下（如《黑客帝国》里的"子弹时间"）那样来辨认数字。然而，他们对流逝的时间本身的感知确实受到很大影响。我们要求他们用秒表记录下他们感觉到的下落时间。（"在心里重建自由落体过程。当你感觉到被释放时按下秒表，然后当你觉得落地时再次按下秒表。"）在这里，与传闻一样，他们对自己的下落时间估计要比对别人的下落时间估计平均延长近1/3。

我们如何理解在自由落体中受试者的时间感觉延长了但他们辨认图像的速率却没有提高这样一个事实呢？答案是，时间和记忆是紧密联系在一起的。在危急时刻，大脑中一个被称为杏仁核的核桃大小的区域转入高速运转状态，征用了大脑的其他资源，迫使一切注意力都转向眼前的情况。当这个杏仁核启动后，记忆被保存在二级存储系统内，由此形成了创伤后应激障碍症日后的闪光灯泡式的记忆。因此，当你处于危急之中时，你的大脑会以更好的"固化"方式来保存记忆。一经重新回忆，这些高密度数据展开来就会使事件显得较长。这也许可以解释为什么时间的流逝似乎随着年龄的增大而加快：你发展出了更高压缩比的事件记录方式，因此要读取记忆相应地就变得较为困难。当你是一个孩子时，一切都那么新颖，丰富的记忆让人感觉到时间的推移是那么缓慢——譬如当你回忆童年的夏天是如何结束时就是这种感觉。

为了进一步了解大脑是如何建立时间知觉的，我们必须了解大脑中的信号是在哪里、在何时建立起来的。人们早已认识到，神经系统研究面临着特征绑定（feature-binding）的挑战——就是说，我们是如何从知觉统一性上来把握对象特征的。譬如说，红色和正方形不是一个流动的红色正方形。从当代对哺乳动物大脑的观点上看，特征绑定通常被正确执行这一点并不令人惊讶，因为在哺乳动物的大脑里，不同类型的信息是由不同的神经流来处理的。特征绑定需要协作——不仅是不同感官

（视觉、听觉、触觉等）之间需要协调，而且同一个感觉通道内的不同特征之间（譬如视觉上就有颜色、运动、边、角等特征）也需要协调。

但脑科学必须解决的问题中还有一个更深层次的挑战，离开它特征绑定几乎是不可能的。这就是时间绑定（temporal-binding）的问题：我们对世界上发生的事情是如何正确分配时间的。这一挑战是，不同的刺激特征是如何以通过不同的处理流运动并以不同的速度进行处理的。大脑必须考虑信息在各感觉通道中以及在不同感觉通道之间的传递速度上的差距，如果它要确定世界上各种事物之间的时序关系的话。

神经信号传递的特征时间尺度有很宽的范围，这一点颇显神秘。人类在进行时间判断时，有相当精确的分辨本领。视觉能够识别的两个刺激可以精确到5毫秒，对它们的时序判断可以精确到20毫秒。大脑是如何以如此精确的分辨本领来识别空间和时间上混乱的信号的呢？

要回答这个问题，我们必须研究一下视觉系统任务和资源。作为其任务之一，视觉系统——位于头盖骨后部的幽暗区域——必须获得正确的外部事件的时间特性。但这个过程涉及提供资源的器官——眼睛和丘脑——的独特性。将信号馈入视觉皮质的这些结构有自身的进化史和特异反应通路。因此，信号会从视觉系统的第一级（例如基于物体的明暗）及时地传播开来。

因此，如果大脑视觉系统要得到正确的事件时序，它只能有一个选择：等待最慢信息的到来。要做到这一点，它必须等待大约1/10秒。在电视广播发展的早期阶段，工程师们曾担心如何保持音频信号和视频信号的同步问题。后来他们意外地发现，他们可以有100毫秒的下降沿：只要信号在这个窗口内到达，观众的大脑会自动同步信号；而超过了1/10秒，声像同步性就很差了，如同观看一部配音低劣的电影。

这个短暂的等待期使视觉系统可以对前期输入的各种信号延时进行平衡。但是它对于追忆过去的感觉不是很有利。动物进化出一种明显的生存优势，那就是尽可能在接近目前的形势下行动。动物都不想生活在遥远的过去。因此，1/10秒窗口是大脑更高层次运行所允许的最小延迟，它可能觉得视觉系统第一级引起的这个时间窗内的延误还在目前这个边界范围内。这个延迟窗意味着，我们的意识是后叙事性的，是一种从事后时间窗口里取来的数据，它只能提供对所发生事件的追溯性解

释。[44]

这种等待最慢信息的战略有巨大好处，那就是使得对物体的识别与照明条件无关。想象一下，一只斑斓大虎从树丛中扑出向你奔来，身形矫健得像阳光下几道剪影。如果老虎的明暗条纹引起的输入信号需要通过不同时间才能感知，那要识别这只猛兽该有多么困难。那样的话，在你意识到将成为老虎的午餐之前你可能只能将这只老虎感觉成时空中的若干碎片。好在我们的视觉系统已经进化到可以协调好以不同速度流入的信息，它的好处就在于无论什么光线条件下你都能识别这是只老虎。

这种假设——系统等待着收集流过时间窗的信息——不仅适用于视觉，也适用于更广泛的其他感觉系统。虽然我们测得的视觉窗口为1/10秒，但听觉或触觉的这一窗口的宽度可能有所不同。如果我同时触摸你的脚趾和你的鼻子，你会觉得这些触摸是同步的。这令人惊讶，因为信号从你的鼻子到达大脑肯定快于从脚趾到大脑。为什么你不觉得是先触及鼻子呢？是不是你的大脑在等待看看还有什么会出现在脊髓管道，直到它确信从脚趾传上来的较慢的信号已有足够时间传递到大脑？这听上去很奇怪，但却可能是正确的。

情形可能是，一个统一的、多通道感受世界的知觉系统（polysensory perception）需要等待传递最慢的总信息。如果仅从沿四肢传导的时间考虑，这便导出了一个怪异但可检验的联想：在过去，高个子的人可能比矮个子的人活得长。对信号传播的等待的结果是，知觉变成了如同电视节目直播。电视节目其实并非真正的直播，而是有一段小的时间延迟，因为播出的内容需要进行必要的编辑。

等待收集所有信息的假说部分地解决了时间绑定的问题，但不是全部。第二个问题是：如果大脑是以不同的速度收集来自不同区域不同感官的信息，那它是如何确定这些信号达到彼此协调的呢？为了说明这一问题，你不妨在眼前打个响指。你眼睛看到的手指动作和听到的响指声似乎是同时出现的。但事实证明，这一印象是你的大脑辛苦构造出来的。因为不管怎么说你的听觉和视觉处理信息的速度是不同的。短跑运动员起跑用发令枪而不是闪光灯就是因为你对声响的反应比对闪光更快。早在1880年人们就已经知道这一事实。生理学在近几十年里证实：听觉皮质细胞对声音的响应要比视觉皮质细胞对闪光的反应更迅速。

这个故事说到这里好像该告一段落了。然而，当我们离开运动反应领域，进入了知觉领域（即描述你看到和听到的东西）后，情况变得更扑朔迷离了。一旦涉及意识，大脑就有一大堆问题需要处理，这些问题涉及如何将外部世界的同步输入信号在感知上取得同步。只有这样，你才会同时感受到发令枪响及其闪光。（发令枪距离你最多30米远，超过这个距离，光速与声速的差异就会造成视觉和听觉接收到的信号不同步。）

但既然大脑在不同时间收到信号，它又怎么能知道外部世界的这些信号应该是同步的呢？它是如何知道枪响不是发生在闪光前的呢？实际上，大脑始终处在对预期的到达时间进行不断重新校准的状态。它是通过一个单一的简单假设开始来做到这一点的：如果它发出一个运动行为指令（譬如拍手），所有的反馈接收系统假定都是同步的，所有延迟都被调整到感知为同时。换句话说，预测输入信号的预期相对时长的最佳方法是与世界互动：你每踢一脚，或触摸或轻敲某件东西，你的大脑就作出一次假设：这些声音、影像和触感都是同时发生的。

虽然这是一种正常的适应机制，但我们已从中发现了一个奇怪的结果：假设你每次按下键，都会引起闪光。想象一下，假如我们偷偷地在按键和随后的闪光之间设置了一段小的延迟（例如200毫秒）。这么小的额外延迟你是意识不到的。但现在如果我们突然撤去这个延迟，你会认为闪光出现在你按键之前，一种动作及其感觉颠倒的错觉。当然，大脑会很快让你意识到这一点，因为它会及时调整延时。

注意，主观时间节奏的重新调整不是大脑在耍把戏，它对于解决因果关系问题非常重要。说到底，因果关系是借助于时序来判断的：我的动作是出现在来感觉信号之前还是之后？解决这个问题的唯一办法是使各种信号的预期时间保持良好的校准，使得即使是面临速度不同的感觉途径大脑也能准确地确定“之前”和“之后”。

必须强调的是，我在这里讨论的一切都是关于自觉意识的。前意识反应似乎显示，运动系统不会等到所有信息都到达之后才作出决定，而是在意识参与之前由皮质下快速通路尽可能迅速地作出反应。这就提出了一个问题：知觉采用的是什么途径？特别是考虑到它滞后于现实，这个问题显得更加突出。是回溯性的，还是通常由自动（无意识）系统从速进行的？答案最可能是，知觉是认知系统能够按后一方式进行的信息

表述。因此，对大脑来说重要的是，需要有足够的时间按最佳解释来说明刚刚发生了什么，而不是拘泥于最初的快速解释。它要将认知系统得来的信息精心提炼成一幅完整图像，就需要花上时间。

神经科医师能够诊断出各种大脑受损的方式，将脆弱的知觉镜像破碎成意想不到的碎片。但现代神经科学一直没问过这样的问题：时间失调该是什么样子？我们大致可以想象，这种病证就像失去色觉，或听觉，或命名能力的病证。但时间构建系统受损是一种什么样的感觉呢？

最近，一些神经科学家开始认为，有些障碍——例如语言生产或阅读——可能起因于潜在的时间节奏问题而不是语言能力本身的障碍。举例来说，患有语言障碍的中风病人在区分不同时间间隔方面更困难，诵读困难病症的阅读障碍可能与无法准确协调好听觉表示和视觉表示之间的节奏有关。[45] 最近我们发现，时序判断缺陷可能是早发性痴呆精神分裂症的症状背后的原因。这些症状有：错误归因（“我的手在动，但我并没有让它动”）；幻听，即正常的内心独白的产生和听觉之间的顺序颠倒。

随着脑时间研究向前推进，我们很可能会发现许多与临床神经学有关的问题。目前，大部分可以想象的时间病症被归类到痴呆症或定向障碍，因为包罗万象的诊断技术错过了许多重要细节。我们希望在未来几年这方面能够受到关注。

最后，对时间问题的远期研究可能会改变我们关于其他领域（如物理学等）的观点。我们的目前的大部分理论框架都包含了牛顿力学意义上的时间变量t，这是一种单向的时间。但当我们开始明白，时间其实是大脑的构造，像色觉一样也会产生错觉时，我们也许最终会从方程中去除这种知觉偏见。我们的物理学理论主要是建立在我们对世界的感知这个滤器之上的，而时间可能是这一征途中最难剔除的。

第14章
超越大脑：为什么智人要从树上下来，为什么其他动物不跟着学？

◎万内萨·伍兹和布赖恩·黑尔

万内萨·伍兹（Vanessa Woods）

《猴有猴的生活》（It's Every Monkey for Themselves ）一书的作者，获奖记者。她在新南威尔士大学获得生物学和英语双学士学位；2004年，获澳大利亚国立大学公众科学意识研究中心科学传播方向硕士学位。她是类人动物心理学研究组的成员，研究非洲倭黑猩猩和黑猩猩心理学。

布赖恩·黑尔（Brian Hare）

杜克大学生物人类学和解剖学系的人类学家和助理教授。他曾在埃默里大学研究人类学、心理学以及人类与自然生态的关系；2004年获哈佛大学博士学位。他的研究以人类认知进化为中心，研究经历包括在西伯利亚狐狸养殖场的工作，在乌干达丛林跟踪黑猩猩，在刚果民主共和国救助倭黑猩猩孤儿。

米克诺坐在那里，右手托着下巴，活像罗丹的雕塑《思想者》。他的左手臂搭在膝盖上，眼神有些模糊，仿佛在沉思。他的黑发仔细地从中间分开，玫瑰色嘴唇红润可爱，米克诺看上去就像人类。但他不是。米克诺是一个侏儒——洛拉亚（Lola Ya）倭黑猩猩的居民。洛拉亚是刚果民主共和国境内的一个非洲大猩猩保护区，专门收养从非洲野生动物肉类贸易中截获的失怙大猩猩。

倭黑猩猩的DNA与我们人类相同的比例（98.7%）要超过它们与大猩猩的比例——因此在他乌黑发亮的头发下，米克诺有一副年轻运动员

的身材，发达的二头肌，6块腹肌也清晰可见。现在的问题是：在它的基因组的30亿个核苷酸中，使得米克诺成为倭黑猩猩而不是一个人的那1.3%的差异在哪里？

几千年来我们一直在设法定义人类的本性。柏拉图将人描述成没有羽毛的靠两腿行走的生物；作为响应，在柏拉图的一次讲座上，第欧根尼抛出了一只拔了毛的鸡。其他定义亦来去匆匆：只有人类使用工具。只有人类会故意彼此残杀。只有人类是有灵魂的。这些个定义就像沙漠中的海市蜃楼，总是飘忽不定。

600万年前，原始人从我们与黑猩猩和倭黑猩猩的共同祖先那里分离出来。从那以后，我们的大脑所发生的变化使我们成为大自然中具有主宰地位的合作者，我们迅速积累起知识，并通过操作工具殖民到这个星球的几乎每一个角落。在进化过程中，我们的进步迅猛而无情。是什么让我们从树上下来，我们为什么会下来？

你认为我在想什么？

当孩子长到4岁，他们开始想知道其他人在想什么。例如，如果你向一个4岁的孩子出示一包口香糖，问她里面有什么，她会说：“口香糖。”可你打开包，结果她看到里面是一支铅笔而不是口香糖。这时如果你对她说，如果我们将这支铅笔再包起来，问等在外面的妈妈，让她猜猜里面是什么，你认为她会怎么说？孩子会回答说：“口香糖”，因为她知道她妈妈并没有看到里面包的铅笔。但4岁以下的儿童一般会说，他们的妈妈会认为包着的是铅笔——因为这个年龄段的孩子还不能摆脱现实世界的诱导。他们认为每个人都知道他们所知道的，因为他们还不能从别人的角度去设想。在这种情况下，他们会认为，人要看到一件东西才能知道它。这种设想别人是怎么想的能力被称为具有某种思想。

人类总是想知道别人在想什么：他刚才看到我看了他一眼吗？那个漂亮女人是不是想接近我？我老板是否知道我不在我的办公室？完整的理论性思想可以解释复杂的社会行为，如军事战略；并据以成立各种机构，如政府。

在整个20世纪90年代，科学家们进行了大量开拓性实验，试图确定黑猩猩——如DNA与我们有98.7%相同的倭黑猩猩——是否具有成熟的

思想。路易斯安那大学（位于拉法耶特）的丹尼尔·波维内利（Daniel Povinelli）让黑猩猩用可看见的姿势向其他人求取食物。这里被求取者或被蒙上双眼，或被套上头套，或用手捂着眼睛，或能够看到它们。黑猩猩一视同仁，它们向那些显然看不到它们的人做出乞讨姿态的次数与向那些能直接看到它们的人做出乞讨姿态的次数一样多。这一研究结果暗示，黑猩猩不具有思想。如果真的如此，那可能就是人类区别于其他动物的关键所在。

在这之后，布赖恩和他的两位同事，何塞普·考尔（Josep Call）和迈克尔·托马塞洛（Michael Tomasello），开始在莱比锡动物园的沃尔夫冈·科勒灵长类动物研究中心着手对雌性黑猩猩雅哈尕（Jahaga）展开观察。实验是这样的：在该中心的一个房间里，你坐在有机玻璃台面后面，台面上有一个放着香蕉的托盘。雅哈尕看得见香蕉，它也可以看到你在观察，并知道如果你看到它走过来，你会拿走托盘不让它接触食物，因为你已经对它这么做过。因此，雅哈尕不是简单地冲着香蕉跑去，而是漫不经心地走到了房间的后面，好像它并不在意你那个不起眼的香蕉，好像它对整个把戏感到厌倦。它继续沿着后墙挪着步子，鬼鬼祟祟地绕到隔板后面直到你看不见它。然后，当它确认隔板挡住了你的视线之后，它便匍匐着快速穿过隔板从托盘里取走香蕉。

这是调查黑猩猩是否会主动行骗（基于他人看不看得见待取目标）的第一个实验。欺骗观察是检验你是否拥有成熟思想的一种重要测试方法，因为在许多情形下，为了欺骗别人你必须知道他们可能会怎么想，然后尝试着操纵局面使他人的想法变得对你有利。在这个实验中，雅哈尕的行为——以及后来各种实验中其他黑猩猩的行为——之所以具有欺骗性，不只是因为它鬼鬼祟祟地绕到一个它知道你看不到它的地方（即是说，它对你此刻所想很敏感），而且还因为它表现出对被欺骗有所察觉：它看上去仿佛是对香蕉没有兴趣（就是说，它会设法让你知道它的这个态度）。

在雅哈尕之后，人们又进行了一系列实验。这些实验表明，在很大程度上，黑猩猩确实会琢磨别人是怎么想的。黑猩猩家族里等级较低的成员总是去寻找首领视线外的食物，因为它们知道首领还没发现它。如果你突然抬头，黑猩猩会顺着你的目光看去，它想知道你在看什么。如果你拖延着不给黑猩猩食物，无论是故意耍它们还是偶尔忘了喂食，那么它们在知道你是故意这么做时所表现出的失望要比知道你只是不经意

失误时感到的失望大得多。但是，这是否意味着黑猩猩就具有与我们人类同样的成熟思想呢？

指向

尽管雅哈尕和其他黑猩猩在一个层面上表现出复杂的心理过程，但在另一个层面上它们的表现却令人失望。如果你将香蕉藏在两个杯子的其中一个里面，而雅哈尕没看到你选择的是哪一个杯子，然后你指着有香蕉的杯子提示它，但雅哈尕却不能按你的手势找到它。尽管你可以轻轻地敲击杯子，或在它上面放置一块色彩鲜艳的积木，甚至围着它手舞足蹈，但雅哈尕选择正确杯子的次数不会比挑选错误杯子的次数多多少。在后来的几十项实验里，它可能会开始猜测这种模式，但是如果你改变提示的方式，譬如说，改为敲击，它仍不能意识到新的线索会帮助它寻找食物。它必须从头学着运用你的新手势。

然而，人类两岁以下的儿童就能够按你的指向找到食物。甚至你只是看一眼正确的杯子，孩子都会顺着你的目光，得到你希望他得到的信息。他们明白你正通过交流隐藏物的位置信息帮助他们。

从黑猩猩的这类实验里我们似乎可以合理地得出结论：运用交际性肢体语言是我们这个物种在进化过程中与其他类人猿分开的一个重要特征。也许这种分享信息的方式使得早期人类发展出比其他动物身上能够看到的更复杂的繁衍形式。但如果确是这样，那么这种能力在开始时是如何进化的呢？

取物

奥利奥是任何孩子都喜欢的那种最好的狗。它会送你到朋友家，然后坐在外面直到你骑自行车载着它回家。它允许你随便拥抱它，此时你的年龄除了拥抱小狗还没有酷到想要拥抱其他人。最重要的是，奥利奥喜欢玩取物游戏。它可以玩到你抬不起胳膊来，因为它可以很容易地一次衔回三个网球。问题是它通常无法搞清楚所有扔出去的球都在哪儿，在捡回前两个后就不知道第三个球扔哪儿了。经过一阵疯狂的寻找后，它只好跑回来看着你，喘着气等待。如果你指出正确的方向，几秒钟后它就会衔回来沾着口水的所有三个球，并准备着你再扔出去。

玩狗的人都知道，当它们想取回什么东西而且它们知道你知道这东西在什么地方，那么它们会像鹰一样注意你肢体语言的任何蛛丝马迹。果然，当布赖恩和他的同事多次用狗进行上述杯子游戏时，只要他们指一指方向，或看一眼，或用脚趾踢一踢藏食物的杯子，小狗立即就能找到藏匿地点（这不是因为它们有灵敏的鼻子——实验中狗在没有视觉提示前不可能确定哪个杯子里藏着食物）。

为什么像狗这样的动物能取得成功，而我们最亲近的亲属却会失败？

一种观点是，狗和我们生活在一起，因此经过数千小时与我们的互动，它们学会了洞察我们的肢体语言。另一种观点是，狼具有集体生活和协作狩猎的习性，而所有的狗都是由这种犬科动物进化来的，狼的这种习性使得包括狗在内的所有犬科动物更适于过一种社会协同的生活。

为了检验第一种观点，你需要和小狗一起玩耍。如果9周大的小狗能够通过杯测试，那么读懂人类的肢体语言就可能不是狗随着年龄增长习得的，而是与生俱来的。布赖恩和他的同事发现，这么大的小狗通过了测试，但仍有一个它们长到9周后就足以能够读懂人类交际肢体语言的问题。为此，它们用饲养在犬舍里很少接触到人类的小狗来测试，结果这种小狗也通过了。

至于第二种观点，你需要花一段时间与大坏狼共处。布赖恩及其同事在野狼保护区对狼进行了测试，并将结果与宠物狗组的结果进行了比对，结果很明显，狼并不比黑猩猩更善于按照人类的社会性提示采取行动。这样看来，狗在过去的4万年里确实进化到可以根据人类的社会性提示采取行动——也就是说，它们经过驯化已经与其狼的祖先分化开来。这一结论的隐含意义着实令人兴奋：在狗与我们人类的长期交往中，作为人类繁衍、合作和语言（人类精神生活的先导和组成部分）的重要发展基础的社会性技能可能已经在狗的进化中得到体现。难道驯化真的可导致这种解决问题的能力发生变化？尽管这似乎是真的，但要检验这种想法，你得去西伯利亚中部。

聪明的狐狸

夏季从莫斯科坐火车到新西伯利亚需要两天的车程。一路上满眼是

缀满鲜花的绿色草原。车到新西伯利亚后，你还得坐上一个半小时左右的车才能赶到阿卡杰姆戈罗多克，这里是当代遗传学最大的实验基地之一。

德米特里·别利亚耶夫（Dmitri Belyaev）被莫斯科的研究实验室解雇了，因为他的孟德尔遗传学与显贵的苏联科学家特罗菲姆·李森科（Trofim Lysenko）的理论相冲突。别利亚耶夫是幸运的，因为对他的处罚随着他失去了莫斯科的工作而结束。在斯大林时代，对李森科的环境获得性遗传理论持有异议是违法的，许多著名科学家因此死在了古拉格[46]。1958年，别利亚耶夫来到新西伯利亚，成为这里的细胞学和遗传学研究所所长。第二年，他在一项孟德尔实验中开始繁殖130只银狐。他用一种简单的方法将一组银狐置于严苛的选择压力之下：那些亲近实验员的银狐可以活着繁殖下一代；而那些对实验员咆哮或对人表现出敌意的银狐则被杀了做成毛皮大衣。另一组，即对照组，则采取随意饲养而不问其对人的态度如何。

仅过了40代后，选定的狐狸开始显露出变化，即你（和达尔文）可能认为需要经过数百万年的进化才出现的那种变化。正如所料，它们变得对人类友好得难以置信。每当它们看到人，它们就会发出叫声，摇起尾巴，嗅着人的身体，舔着人的脸。但更奇怪的是它们生理上的变化，这种变化比对照组发生的频率高很多。选定的狐狸的耳朵变得耷拉，尾巴变得卷曲。它们的皮毛失去了伪装，变得参差不齐；前额出现了星状纹路，头骨也变小。总之，它们的模样和行为非常接近其近亲——家养的狗。

现在，我们开始进行一项大的检验。如果狗能够在驯化过程中获得社会技能，那么选出的银狐或许也能获得这些技能。

事实确实如此。驯养的银狐可以像狗一样读懂人体语言。而对照组则不能。

银狐对人类的社会性暗示的洞察能力是整个谜团的关键一环。人们（包括作者）原以为，在狗身上发现的非同寻常的社会技能可能是进化的结果，因为在驯化过程中聪明的狗更容易生存和繁殖。但是，别利亚耶夫的狐狸并没有繁衍出比一般狐狸更聪明的品种，只是更友善而已。由此看来，选定的狐狸更善于洞察人的暗示只是一种失去了对人恐惧的

副产品，这种对人的恐惧已替代为与我们的密切互动。狗的社交技巧很可能在其驯化过程中也是经过类似的过程。为了捡拾人类居住区附近的垃圾，原初的狗已经失去了我们的恐惧。后来，出于偶然，在与我们的交往过程中，它们开始采取原本只是在它们之间运用的社交技能——就像我们只是其群落的一部分。

最重要（也是最有争议）的是，人类在进化中可能也经历过类似的过程。能够生存并繁衍下一代的可能并非如人们经常暗示的那样是最聪明的原始人，而是更善于交际的原始人，因为他们在共同解决问题方面表现得更好。善交际的原始人获得了更高的适应性，并会随着时间的推移来考虑选择更复杂的问题来解决。人类变得聪明只是因为我们首先变得更友善。

黑猩猩的缺陷

合作是人类成就的基石。从某种程度上说，这有赖于我们有先进成熟的思想和对社会性暗示的运用。但人不是唯一的熟练掌握合作技能的物种。是人的什么品质使我们成为如此灵活的合作者？或者换一种说法：黑猩猩在合作方面到底出了什么毛病？它们同样过着高度社会性的群居生活，一起觅食，群落内保持着等级关系。那么是什么妨碍了它们在通过合作与交流来解决问题方面没能变成像人（或狗）那样灵活？

恩甘巴岛黑猩猩保护区是乌干达境内位于维多利亚湖中心的一片延绵数百英亩的原始森林。在晴朗的日子里，你可以听到一水之隔的黑猩猩的喘气声。现在，两只黑猩猩，基多格和康妮，正面临着一种两难境地：木板两端堆满了食物，可木板正好在可以够到的距离之外。为了将木板拉近点，它们俩必须拉动穿过木板上金属环的绳子。如果只有一个人拉，那么绳子只是车轱辘转，木板并不动。在此情形下，基多格作为女首领，会一把把康妮推了过去，它来拉康妮的绳子，于是只听见滑轮吱吱叫，谁也没得到食物。

这种行为很令人费解，因为野生黑猩猩都是伟大的合作者，经常以一种复杂的组织有序的方式猎食。但也许这种合作的背后并没有多少思想基础，它可以简单地归结为每个动物都希望得到同样的东西，因此在同一时间里能够一起工作。成功只是意外，合作只是一种表象。

但是，如果你在喂食的时候观察基多格和康妮，你会发现它们并不是分享食物。如果康妮有一块食物而基多格待在一旁，那么基多格多半会偷它的食物。另一方面，在保护区里一起长大的萨莉和贝基则像姐妹一样，它们在任何时候都能够和平地分享食物。当你给它们做绳子测试时，它们第一次就成功了。

显然，如果考虑到容忍这一点，黑猩猩是可以自发合作的。它们不仅知道什么时候需要有人帮忙，同样也知道谁是好的合作伙伴。马瓦是另一个黑猩猩首领，它就不是一个好的合作伙伴。它等不及它的合作伙伴拿起绳子的另一端就一个劲儿地拉，木板自然一点不动。另一方面，布万巴莱则是一个伟大的合作者，它等待合作者，它们几乎总是能成功地获得食物。起初，其他恩甘巴黑猩猩平等地选择马瓦或布万巴莱做合作者，但在经历了马瓦的拙劣表现后，大多数黑猩猩下次都选择布万巴莱做合作者。

然而，黑猩猩的这种合作是极为有限的。黑猩猩一般只与熟悉的成员合作，与它们分享食物。如果它们不认识或不喜欢那个潜在的合作伙伴，那么无论有多少食物在等待它们享用，它们也不会合作。但人类则不同，人是靠合作生存的，即使要合作的那个人你可能不认识，在大多情形下更谈不上喜欢。（你有老板吧？）这种高度的社会容忍度可能是人类独有的合作形式的基石之一。

因此，缺乏容忍也许是黑猩猩无法发展出更加灵活的合作技能的主要障碍之一。但人类还有另一个近亲，一个通常被我们遗忘了但可能比我们预想的更喜欢我们。

失散已久的表亲

黑猩猩家族是一种由雄性主导的社会，杀婴行为和其他形式的致命侵犯行为不时发生。与此情形相反，倭黑猩猩则生活在一种非常宽容平和的、由雌性主导的社会里，雌性首领通过性行为维护着团体的凝聚力和必要的张力。

既然倭黑猩猩比黑猩猩宽容，那么这对它们的合作能力意味着什么呢？

人们对恩甘巴岛的黑猩猩进行了进一步的实验。只要食物是分两处放置，或放在木板的两端，大部分黑猩猩都能够合作得很好。但只要你把食物放在板的中间作一堆放置，黑猩猩间的合作便土崩瓦解。即使参与测试的一些黑猩猩相互间相对宽容，并且此前已多次通过绳/板测试，但只要食物可由黑猩猩首领垄断，其他黑猩猩通常就会拒绝拉绳。

当我们对倭黑猩猩做同样的测试时，它们通过游戏和性行为进行彼此商谈——尽管这是它们的第一次测试。倭黑猩猩都以性能力强健而闻名。雌性会在一起揉搓阴蒂，雄性之间也会有性交活动，无论年龄或性别都是如此。性生活既是群体里舒缓紧张的一种方式，也被用来平复暴躁的脾气，或结成联盟。群体里还可见协商活动，这使得倭黑猩猩具有很高的容忍度。

因此，我们得出的结论是：黑猩猩之间有合作，但缺乏很强的包容性；倭黑猩猩非常宽容，但在野外并不真正进行合作。原始人之所以能够在600万年前从我们与黑猩猩和倭黑猩猩共同的祖先那里分化出来，靠得可能就是我们变得非常宽容，这使我们能够以全新的方式合作。如果没有这种高度的容忍性，我们不会成为今天这样的物种。

在非洲寻找我们的头脑

自发合作并非倭黑猩猩比黑猩猩更像人类的唯一特征。像人类一样，倭黑猩猩的性别差异不是那么明显。雄性的身体并不比雌性的大很多。雌倭黑猩猩像人类女性一样，起着强有力的纽带关系，而雌黑猩猩一般则不是这样。人类和倭黑猩猩有类似的脾气，因为我们都会规避风险并对新来者满怀戒心。

理解倭黑猩猩对理解我们人类之所以为人至关重要。不幸的是，它们的数量正在快速下降。刚果民主共和国是它们唯一的家园，不断爆发的战争使得对它们的研究变得困难。非洲的类人猿保护区，包括刚果民主共和国境内的洛拉亚倭黑猩猩保护区、恩甘巴岛和钦穆蓬加（Tchimpounga）的黑猩猩保护区，为我们提供了一个令人兴奋的深入了解我们的近亲的思维的机会。与人工饲养下患有慢性心理和生理疾病的实验室动物不同，保护区的猩猩以很大的社会群体生活在热带雨林地区。半放养的猩猩可以像传统实验室进行的那样在室内进行实验观察，但成本要低得多。保护区的动物并没有异常行为（例如摇晃或吃粪

便）。初步数据表明，它们比被囚禁的猿类更胜任各种身体任务，因为它们的日常环境非常丰富。

坐着像罗丹雕塑的倭黑猩猩米克诺于2006年9月死亡。尸检报告显示了它的脑部有挫伤，这表明它可能因从树上掉下来摔成脑震荡而死亡。米克诺的密友伊西罗坐在它身边不愿意离开，它能理解什么是死吗？它是否会像人一样感到悲伤？在寻求我们之所以为人的漫长道路上，我们仍然有很长的路要走。但即使我们在不懈努力，要搞清黑猩猩之所以成为黑猩猩，倭黑猩猩之所以成为倭黑猩猩，我们仍然有成千上万的问题有待回答。

第15章 我们当中的外来者

◎内森·沃尔夫

内森·沃尔夫（Nathan Wolfe）

斯坦福大学人类生物学Lorry I.Lokey访问教授。他于1998年获哈佛大学免疫学和传染病学博士学位，先后荣获富布赖特奖学金（1997），国立卫生研究院主任先锋奖（2005）和美国国家地理新兴浏览器奖（the National Geographic Emerging Explorer Award, 2009）。沃尔夫的研究方向是结合分子病毒学、生态学、进化生物学和人类学等方法来描绘地球上微生物的生命多样性。他的主要成果是第一次给出了反转录病毒从非人灵长类动物到人的自然传播的证据。他发起并指导全球病毒预测行动（GVFI），这是一套监测新传染源从动物向人类蔓延的早期预警系统。GVFI最近从Google.org和斯克尔基金会获得1100万美元的资助。它协调着来自世界各地的100多位科学家和工作人员的研究活动。目前已在喀麦隆、中国、中非共和国、刚果民主共和国、刚果共和国、加蓬、赤道几内亚、老挝、马达加斯加、马来西亚、圣多美和普林西比等国家展开积极研究并设立公共卫生项目。

想象你正坐在家里舒适的沙发上，手里端着咖啡杯，腿上搁着书，孩子的游戏声从另一个房间传过来。一切似乎都很正常——就像它一直都是这样。但是话说回来，事实也许并非如此。也许，你隐隐约约地感到你不是一个人待着。

突然，面纱揭开了。你周围到处是看不见的生命群落——沙发上，咖啡杯的盖子上，甚至包括你腿上的书的封套上。一种无法识别的生物，虽小但明显是活的，它没有细胞，没有酶，并且显然缺乏我们所熟悉的生命体应具有的大部分生物机制。这些外来者无处不在，它毫不夸张地将自己整合到我们周围的生命组织中，整合到组成我们这个世界的

每一种细菌，每一种植物，真菌，动物体内。它们已入侵了，并且也许已经赢了。这就是……病毒！

对外来者的害怕总是纠缠着人类——通过电影、书籍和孩子的梦。它激发起我们对太阳系行星的好奇心——火星上有过水吗？火星上有生命吗？它们是否有大气？是否有氧气？它驱使我们对我们的宇宙进行探索，推动我们开发出巨大的射电望远镜阵列来搜索宇宙中遥远的智慧生命的呐喊。

然而，具有讽刺意味的是，我们在搜索广袤的宇宙时，很大程度上却忽视了一颗特别重要的行星。它具有维持生命的所有要素，而且探索起来相对容易——根本无需航天服或飞船。这颗星球上有我们已经知道的生命形式，而且很可能还存在我们尚未知道的生命形式。它就是我们的地球。当我们在宇宙中寻找外星人时，我们不幸地忽视了它。

我们已经知道，地球上的生命有多种类型。地球不仅供养细胞生命（细菌及其最近才被认识的、长得十分相似的堂兄弟——古菌）和真核细胞生命——真菌、植物和动物，而且还供养另一种颇不寻常的生命形式——朊病毒。这项发现荣获了1997年度的诺贝尔奖。朊病毒是一种奇怪的生命形式，它不仅没有细胞，而且缺少DNA或RNA这样的地球上所有其他已知生命形式用作模板的遗传物质。然而，朊病毒顽强地存活下来并且似乎可复制，使其他生物染上疯牛病。

我们的地球还供养着病毒——一种没有细胞的微小生物体。病毒可以惊人的效率侵入地球上其他生命体，是一种裹着蛋白质外套的纯粹的遗传物质，它们依赖于宿主细胞，离开宿主它们既不能生长，也不能复制。

与已有的关于病毒的文章不同的是，本文不打算着重写病毒对人类的伤害，而是跳出地球外来者看待我们这个星球上病毒的多样性和丰富性，其生态意义以及它对提高细胞生命体丰富的生命形式的重要作用。可以说，对病毒更为深刻的理解为我们提供了新的观点——不只是针对人类健康和疾病，而是对我们这个星球上的基本生物学的认识。此外，由于病毒具有如此特异的生命形式，了解它们的繁盛可能会使我们得到关于真正外来者的线索。

它们无处不在

地球是一个活的星球。即使是从太空上看，它也展示出丰富的生命印迹。然而，我们从太空上看到的只是全部真相的一部分，从生物量和多样性来衡量，我们这个星球上生命的主要形式是微观的。微观的生命形式无处不在：海洋里、陆地上和地下深处。利用深海勘探技术在海底钻探几百米深后取得的岩样显示，这些地层中存在生物体。它们不是依靠太阳能而是依靠地心产生的热量存活。虽然细胞微生物，如细菌和古菌，也许更容易发现，但以病毒形式存在的无细胞微生物则代表了这种多样性的重要组成部分。

据认为，每一种宿主细胞生命形式体内至少有一种类型的病毒。每一种藻类、细菌、植物、昆虫、哺乳动物，无不如此。即使每一种细胞生命体只容留一种独特的病毒，这也将使病毒成为地球上最多样化的已知生命形式。而实际上，很多物种，包括人类，寄宿着大量的各种病毒。

虽然病毒是又小又轻——已知最大的阿米巴感染性拟态病毒也只有600纳米长——但它们留下的生物学足迹却可观得令人难以置信。在1989年发表的一篇具有里程碑意义的文献里，卑尔根大学的厄伊温·贝格（Oivind Bergh）及其同事报告了他们用电子显微镜对未污染的自然水生态系统里感染细菌的病毒进行了计数，发现每毫升水中有多达2.5亿个病毒颗粒。其他更全面的对地球生态系统中细菌类病毒的生物量的统计更是大大超出人们的想象。根据一项估计，如果将病毒头尾相连排列起来，由此得到的长链将长达2亿光年，远远超出我们银河系的边缘（银河系大约只有150光年），其质量估计则相当于约100万头蓝鲸的体重。[47] 病毒的影响岂止是一点点，它们简直覆盖了我们的星球，其丰富性和影响力我们才刚刚开始了解。

朋友还是敌人？

如果你像大多数人一样，每当想起病毒，脑子里就会浮现出一种具体念头：一种致病的微小的病原体。如果你有机会开启一个能永久消除地球上所有病毒的开关，你也许就会选择打开开关。但你仔细想过没有，寄宿在地球上几乎每一种细胞有机体内的所有这些病毒，有着如此

惊人的多样性和丰富性，难道就只有负面作用？答案是：绝对不会！

尽管病毒必须感染生命体的细胞才能完成其生命周期，但这并不意味着它们的命运就是引起破坏，我们地球现有的平衡依赖于病毒世界的行动，它们的灭杀作用有着深远的影响。据估计，海洋系统中每天有20%～40%的细菌是由病毒杀死的。这些死亡提供了巨大的有机物质来源，它们向海洋环境释放出氨基酸、碳和氮，而在这种海洋营养物循环中病毒起主要作用。[48] 此外，虽然关于病毒在调节生物多样性方面的作用的研究还处在起步阶段，但人们认为，病毒有助于维持像细菌这样的重要的环境角色的多样性，这里病毒起着“联邦反托拉斯检察官”的作用——就是说，它可以防止任何一种细菌变得占绝对优势。

但是，病毒除了在养分循环和多样性维持方面起着核心作用（当然，这是通过病毒介入对活细胞起破坏作用）外，它对生物个体是否具有某种潜在的益处呢？不幸的是，对人类病毒的大部分研究侧重于病毒的病原体作用，这种负面偏见不仅影响到我们对病毒的认识，也影响到一直以来我们对病毒的研究。因此毫不奇怪，我们熟悉的大都是病毒造成的伤害。对于是否存在有益病毒的问题至今还没有系统的研究。如果我们寻找“好”病毒，我们可能会发现什么呢？

病毒与其宿主以及其他生物体的互动关系是一种连续统：有些病毒损害其宿主，有些则对宿主有好处，还有一些——可能大部分——则与宿主形成一种相对中性的关系，对于它们为了存活至少必须是暂时居住的生物体，既无重大伤害，也谈不上有利。虽然对有益病毒的研究事例不是很多，但一个偶然的发现——黄蜂、毛虫与病毒之间的复杂关系——已经为求证病毒有益于宿主的可能性带来了一线光明。我们通常将黄蜂看成是筑巢生活的群体性生物，但也有许多种黄蜂并不采用群体蜂巢的方式养育幼蜂。其中一种是非集群性茧蜂科黄蜂，它将卵产于其他昆虫的幼虫体内，当蜂卵孵化成幼虫后，幼虫就将所寄生的毛虫作为食物。不用说，这是以牺牲其他幼虫为巨大代价实现的，其结果是在进化上形成装备竞争，毛虫发展出对付黄蜂产卵的防御技术，黄蜂则需要寻找出对策。这其中的一种令人拍案叫绝的反制措施是病毒与宿主的合作，这种病毒叫多态DNA病毒，是一种用DNA而不是RNA作为遗传信息的病毒。这种病毒已进化出一种与黄蜂的互利关系。它在黄蜂的卵巢内复制并同蜂卵一起注入毛虫体内，病毒能够抑制宿主毛虫的免疫系统，从而有利于保护蜂卵。黄蜂帮助了病毒，病毒也帮助了黄蜂。[49]

鉴于病毒无处不在，因此如果只看到它们的破坏作用就未免让人感到惊奇。进一步研究很可能会揭示出这些生物体行为深远的生态重要性：病毒感染不只是具有破坏作用，而且对被感传的许多生物有益。

进化调查者

在进化的核心问题上，病毒可以提供的好处之一是遗传的多样性。由于病毒具有很高的突变率和很强的彼此交换遗传信息的能力，因此成为巨大的遗传变异产生器。它们不断地将遗传信息从一种生物体带到另一种生物体，随机地将自身的和前宿主的新的遗传信息带给新的宿主。这种遗传多样性相当于进化过程中的动力，有利于通过自然选择形成新的基因。虽然病毒携带的基因并不总是有利于新的宿主，但它们提供至关重要的好处的情形也非常多。典型的例子如某些细菌（譬如霍乱）产生的毒素就是由这些病毒基因演化出来的。霍乱毒株携带一种名为CTXf的病毒，这种毒素带有造成大量腹泻的基因，从而帮助霍乱细菌迅速蔓延。

但是，这种基因引入方式绝不限于细菌。反转录病毒是一种能够将其遗传物质整合到宿主DNA的病毒，它可以寄宿于各种动物。现在看来，反转录病毒基因引入到类人猿祖先从而产生新的哺乳动物基因的这种机制可能在我们人类的起源上起着重要作用。[50] 基因ERVWEi显然来自反转录病毒，现已成为人类特有的和永久的基因。由于反转录病毒及其产物必须与免疫系统的复杂性进行交涉，而免疫系统则不断努力地拒绝“外来物”。这也许并不奇怪，反转录病毒基因可能利用我们的生殖系统。据认为，ERVWEi恰恰具有这一功能——就是说，它有助于抑制免疫系统，使得“外来”胎儿免遭拒绝，就好像反转录病毒必须控制免疫系统以避免自己遭到拒绝。通过遗传多样性的产生和运送，病毒为地球上生命的进化提供了重要的创新动力和流动性。

为病毒平反

“病毒”一词来自拉丁语virus，意思是“毒素”或“毒药”。虽然从这个词本身首次进入英语（在14世纪，作为“有毒物质”的同义词引入）以来，我们对病毒对于地球上生命的重要性的理解已经有根本的转变，但病毒科学却还在延续着这些早期的负面的误解。一些病毒学家已经开始

接受对病毒作用的更全面的看法，但广大民众以及众多研究脊椎动物病毒的科学家却还是只盯着那些对人类和动物种群有危害的病毒种类上。

对病毒的重要性的误解也影响到研究这些病毒的学科。病毒科学通常被看作是微生物学和人类医学的分支学科，这种认识必须予以纠正。病毒的研究必须提升到基础科学的地位。对全球范围病毒多样性的刻画和理解，无论是从预防下一个主要疾病传染途径方面说，还是从了解海洋碳循环角度看，都符合人类的近期利益和长远利益。演化生物学、物理学、地球科学和计算机科学领域的新一代思想家应当与医生和分子生物学家联合起来，共同促进我们对我们这个星球上最重要的生命形式的了解。

幸运的是，我们有了让病毒科学曙光初露的新工具。环境基因组研究已开始揭示出我们这个世界上病毒和其他微生物的生命的丰富性。这项研究是利用基因测序技术来刻画从土壤到皮肤等各种丰富的生物环境中基因的多样性。例如，对人类粪便的研究已检测出粪便样本中存在1000多种不同的病毒——病毒种类要比人们认为的整个人类受感染的已知病毒的种类多得多！[51] 显然，我们对居住在我们体内和我们这个世界里的病毒的了解还处于初级阶段。对动物、植物、土壤和水系统中生存的病毒的进一步研究代表着做出发现的新的航程。对病毒的收集将形成一座21世纪的病毒博物馆，科学家可以用它来了解环境，预测下一次疫病的流行，开发新的抗生素，并为各种用途找出和引入新的基因。

一个新的、全方位的病毒科学具有巨大的潜力。病毒已被证明是一种非常宝贵的疫苗和分子生物学工具。让未成年人接种牛痘病毒疫苗（牛痘病毒的变种）使人类彻底消灭了天花，这种疾病曾是人类面临的最严重的疾病。噬菌体（细菌性病毒）在分子生物学里被证明是做出基本发现的重要模型，它也是向活细胞引入新基因的分子工具（载体）——最近非胚胎干细胞的成功创建尤其引人注目。[52] 我们无法想象，离开这些工具我们还能够做出多少基础性的和实用性的发现。

病毒研究除了具有全球病毒综合调查这一实际意义之外，还可以告诉我们有关我们在宇宙中位置等方面的信息。病毒代表着一种独特的、更为简单的生命形式——一种在宇宙中比复杂生命形式更具普遍意义的生命形式。[53] 它们或许可以告诉我们，如何才有可能发现地外文明并与之互动。在地外发现病毒已被证明是一项艰巨的任务，更遑论什

么“第一次接触”了。这很可能就是个错误的假设。我们几乎还谈不上揭开病毒世界的面纱，这该使我们确信，要确认真正的外来者将会遇到令人难以置信的困难。

大众想象中的外来者通常都被描绘成不是心地十分善良就是十足的邪恶。我们对病毒的成见也类似。在现实中，他们代表了全方位的好、坏、丑陋和善良。假设外星生命会有所不同那几乎可以肯定是错误的。另外，所谓外星人是否会注意到我们这种想法基本上是基于宇宙的人类中心论的认识。第一次接触很可能是一种戏剧性遭遇，既可能是与具有较高智慧的善良之辈相遇，也可能是遇到来自另一个星球的邪恶的战争狂人。但更可能的是，它们很难被发现，即使它们生活在我们周围的空间里。因此，不论是从物种上说还是从带给我们科学上巨大益处上考虑，都谈不上非常好或是非常糟糕。

第16章
社群性昆虫是如何变得具有社会性的？

◎塞利安·萨姆娜

塞利安·萨姆娜（Seirian Sumner）

伦敦动物学会动物研究所进化生物学研究员。她1995年获伦敦大学学院动物学学士，1999年获伦敦大学学院行为生态学和进化方向博士学位。随后在哥本哈根大学和巴拿马史密森热带研究所工作，2004年进入动物研究所。她曾荣获多项研究职位，包括最近的为妇女设立的L'Oréal科学会员。

萨姆娜的研究重点是社会性的进化——组织性完好的社会是如何演变来的以及这种社会行为是如何被继承下来的。她一直与世界各地的各种蜜蜂、黄蜂和蚂蚁打交道，通过观测、实验操作和分子水平的分析（包括基因表达）来研究它们的行为。她对社会性的起因和在这种重大进化转变中基因所起的作用特别感兴趣。

你需要搬迁到一个新的地方来容纳不断壮大的家。你肯定想待在一个搬迁成本较低的地方，因为你真的很喜欢当地的杂货店。经过一阵搜寻，你和你的家人决定了新家的住址。在你的指导下，全家组织成一个高效的团队：年轻的帮助收拾老房子，年长的将行李搬到新家。女孩们承担了大部分工作，所幸的是你有很多女儿。女孩也善于组织，因此新房子很快就落成了并进入运作。孩子们兴奋地告诉兄弟姐妹，在来的路上她已发现了糖果店。

世界上任何家庭都可能发生上面的故事，但我们这里指的是一个蚂蚁家族。像蚂蚁这样的社群性昆虫的生活与我们人类的社会生活之间的相似性简直令人震惊，远远超出了搬家的比喻。垃圾清理、上岗值班、照看幼儿、君主制度、奴隶贸易、自由骑手、处罚、叛乱、建筑、农

业、杀戮、同类相食——社群性昆虫可谓一样不落，在我们之前的8000万年前它们就这么做了。

其实，我们对社群性昆虫感到着迷，对它们是如何处理冲突与合作之间的平衡感兴趣不足为奇。一个多世纪以来，科学家们描述它们，观察它们，测量它们，甚至将它们集中起来以研究它们的基因。关于这些庞大帝国是如何演化和维持的方面的知识，我们现在已有相当完备的科学文献。我们知道，这些动物都是非常出色的信息传播者：蜜蜂会用著名的摇摆舞来描述彼此间发现的食物位置，其精确性可以和汽车上的卫星导航系统相媲美。它们是宽容的互助者：切叶蚁与真菌花园之间的那种专有的联盟关系使得最善表现恩爱的好莱坞佳偶都显得老土。它们还是土地的征服者：阿根廷蚁（学名叫Linepithema humile）的一个家庭可以延伸到6000千米远，远远超过任何人类的城市群，更不用说单个的人类家庭了。仅蚂蚁一项的生物量就占到世界动物生物量的20%，远远超过人类的总和。毫无疑问，真社群性昆虫——那些过着四世同堂，有着明确的劳动和合作育雏分工的集群性物种——是大自然的一个奇迹，值得我们予以充分注意。

真社群性昆虫成功的秘诀在于它们的任务划分。个体成员职能专一，要么专事繁殖（女王等级和雄性），要么专事觅食和育雏（工人等级，所有这些都是雌性）。在蜜蜂王国（这是一种发展得最为成熟的社群性昆虫），工蜂已失去了交配能力，因此不能产下受精卵。整个王国的全部基因遗传任务只交给一个个体承担，这个个体就是蜂王。这种适应性一旦得以实现，等级制就会顺利地演化出来，使个体忘却了这种等级的永恒性。进化将粗活儿交给工蜂去处理，因为它们承担着除了交配以外的一切工作。这也是查尔斯·达尔文在创立他的自然选择理论时最感头痛的地方。但现在他可以安息了。现在我们知道，演化出永久不育等级无疑是社群性昆虫可以做到的最聪明的事情。等级制可以减少冲突：如果你不能产卵，你就不会把时间和精力浪费在挑战具有生殖力的领导身上。较少的冲突可以使你的基因得到遗传，只不过是以亲戚（通常是你的兄弟姐妹）后代的形式。通过限定王国里王后的数量，员工们可以更清楚地知道它们这一窝里彼此间的相互关系——由此它们通过工作来回报自身的基因遗传。专制的社会结构一旦形成，这个群体就会迅速成长，成千上万的工蜂（或工蚁）就会心甘情愿地供奉着唯一的女王。

让我们沿着真社群性昆虫的复杂的演化轨迹进一步探索下去。工虫的等级又可以细分为多种形态：大师傅级的、小师傅级的、刚出道的和刚入门的，各个职能都不一样，由它们的个体大小决定，使得整个种群的基因传播链运行得更加顺畅。最后，在我们认为是社会复杂度极限的地方出现了第三个等级：大而凶猛的个体，这些是兵虫，它们的下颌长着尖牙和锋利的长矛，是种群王国的守卫者。这种社会结构曾启发了许多科幻小说作家——他们依据这些事实创作作品。

这些高度复杂的社会是如此令人振奋和易于理解，一下就抓住了社会生物学家的注意力。但它们能够提供给我们的关于真社群性起源方面的信息却不多，因为这种社会结构是进化的结果，我们很难从结果来反推原因。下一代社会生物学家最终的问题是：是什么力量造就了这种社群性？我们想知道是什么条件和选择压力使得社群性昆虫的祖先翻越窗台，演化出成熟的社群性结构。这里至关重要的是，最早的种群是如何演化的？

让我们从最初的情形开始。社群性昆虫的祖先是孤独的，我们很容易在自然界中发现这样的活标本。单独生活的挖掘蜂将卵产在地上挖好的洞里，并捉来多汁的毛虫作为待孵化出来的宝宝的食物，然后就将幼虫遗弃在这个“食品柜”里让它自己孵化成长了，因此挖掘蜂的幼虫打一开始就不知道谁是它们的父母（每个青少年的梦想）。接下来，我们有了最简单的昆虫社会，它由两个或两个以上的共居一巢的雌性组成，它们共同承担任务——筑巢、看家和觅食。如果能证明这样做比单独筑巢更易成功，那么我们就迈出了走向真社群性的第一步。因此社会演化的最终突破是要解决独居生活（如寂寞的挖掘蜂那样）向社群生活的转变问题——这个社群可能很小，如由全能的雌性组成的平等的社群；也可能很大，如由高达百万成员供奉一个王后这样的高度组织化的帝国。事实是，能够发展成真社群性的昆虫物种只有1/50，因此可以肯定，由独居向社群生活的转变是困难的。问题是为何这么难？

自然世界不乏卓越的数学家（动物、植物、真菌、藻类都是），从遗传适应性的演化尺度上看，它们都是平衡各自生活模式的好处与相对成本的高手。已故的演化生物学家W.D.汉密尔顿指出，社群性昆虫要比一般动物更聪明，因为它们需要评估与同巢配偶共享多少基因。如果它们的同巢配偶是近亲的话，那么社群里的其他成员就将接受这种看似不公平的生殖分工。用术语来说，汉密尔顿法则可以表述为：如果帮助亲

属繁殖获得的收益（b）[其大小由你与这些同巢兄弟姐妹的关系远近（r）来衡量]大于自己繁殖的成本（c），那就值得过社群生活——也就是说，如果br>c，昆虫必然选择社群性。再也没有其他的科学表述能比汉密尔顿法则更简单、更优美的了，它使我们能很好地理解真社群性的演变。但它不能解释真社群性阶梯的第一步是如何实现的。最先需要用汉密尔顿法则进行判断的条件是如何产生的？雌性是如何寻找亲属并决定同巢合作的？

取得最终突破的关键被锁定在社会多态性蜜蜂和黄蜂物种上。它们可能就是我们要找的独居生活和社群生活之间的环节。这些物种可以选择是过独居生活还是过社群生活，这主要取决于它们在哪儿生活，在哪儿出生，邻居是谁以及天气怎样。如果我们能够了解这些决定是如何做出的，那么我们就开始了解真社群性昆虫祖先的行为了。为此我们先来了解一下社会多态性蜜蜂的世界。它的困境是什么，它是如何解决这些问题的？

困境1：是否要过社群生活？

我们来考虑这种蜂，在英格兰春天的微风中，一只雌蜂从休眠中苏醒过来。它能够估量出选择群居而不是独居的好处吗？这个决定就如同是否要租一间公寓：你可以选择与其他人合租，那样较便宜也较有效率（你可以分担家务、一起去购物等），但你必须忍受其他人邋遢和不合群所带来的不便。另一种选择是租个单间，但那样较昂贵，也较担风险：你不仅需要一个人操持所有家务，而且出门购物时家里被盗了怎么办？从积极的方面说，至少你有宁静，有私密性，最重要的是有控制权。当我们做这些决定时，我们需要权衡这种选择的成本和好处，就像我们这只刚苏醒过来的蜜蜂。

如果这只蜜蜂出现在苏格兰北部，它迎来的可能是一个清冷而短暂的夏天。它得赶紧独自筑巢，而且仅能繁育一窝后代，这些后代明年就将成为王后。如果这只蜂是出现在英格兰南部，那里夏季长而温暖，你会发现有同一品种的蜜蜂独自筑巢，就像它们的北部同类那样。但你也会发现同一品种的蜜蜂选择群居生活的。在这些真社群王国里，第一窝蜂将成为工蜂，其任务是帮助繁育第二窝蜂，后者才是明年的王后。显然，是选择单过还是群居，环境因素有着重要影响。

在学校里，我与几乎同一群女孩交往了5年。这个小组的层级由自我任命的头儿进行仔细监管，她是通过在众人中轮值“最好的朋友”这样的方式来行使权力的，轮值到谁依她的情绪而定。这是一种分而治之的策略，让每个人都感到自身的脆弱，但又有希望，偶尔还心怀感激。能在一个组织里立住脚是受欢迎的一个标志，而且确保了不乏男孩的追随。我们的蜜蜂之所以愿意过集体生活的一个好的理由就是它认为自己能成为王后，或者至少成为一个受人喜欢的下属（“最好的朋友”）。有些蜂群中的个体间可以是平等的，在这种情况下，作为一个工蜂并不真正意味着什么，只是有些个体可能比较专制。而在顶级社群里，蜜蜂则需要知道，如果自己参与进来，成为王后的机会有多大。在过去的15年里，已经有一大堆理论模型试图解释这种繁殖分配方案是如何决定的。根据这些模型，社群性昆虫王后掌控局面不是像上述学校的女首领那样靠人气来确保组织的凝聚力，而是依社群大小、环境和物理等因素的不同来激励组织成员。就像学校的女首领，社群性昆虫的女王需要有自己的社群，所以它会说服其他雌性留下来，帮助它共同繁育后代。胡蜂似乎就认可这样的社会契约：雌性蜂后允许其下属产卵，为的只是让它们高兴。其他真社群性种群，如allodapine蜂，则会为更好地分享生育机会而相互竞争。

无论采取什么方式，即使在最简单的种群里，社群性昆虫都能够（间接）衡量出不同生活方式的相对成本和获益。在一个所有雌性都有可能成为皇后的简单种群里，生殖权益是如何划分的？我们正在逐渐累积这方面的数据。但我们似乎并没有得到任何确切答案，每一个新种群的研究似乎支持不同的理论模型。也许我们应该放弃寻求普遍法则的努力。也许我们一直是从一个错误的角度在看问题，以一窝或一个种群作为研究对象得到的信息可能更丰富。

困境2：如何成为社群中的一员

当蜜蜂决定过群居生活后，它需要找到一个合适的群体。它到哪里找呢？她怎么能肯定这个对象是近亲？在这方面犯错误——选择一个其中女伴与自己毫无关系的巢来安家——将是灾难性的，正像工蜂，只能通过帮助其他伙伴生养来遗传它们的基因。要认清亲缘关系可以运用如下的简单法则，譬如像“与在你附近冬眠醒来的主儿同巢”，因为你的姐妹往往就在你醒来的附近冬眠。再譬如像“与长得和你很像的主儿同

巢”，因为每个物种都有确定的特征，例如理查德·道金斯说的“绿胡子”标签。他在《自私的基因》一书中指出，这种独一无二的特征是“基因在众多其他个体中‘辨认’自己副本的一种途径”。

我有一个同姓的朋友。我想借此说明的是，这个姓是我们可能有亲缘关系的一个标志，因为我的这个姓不是很常见。而且她也确实看起来有点像我（当然，我们没有绿胡子）。然而，她来自于不同的国家，除非我们去做基因测试，否则我无法确认我们之间的亲缘关系是不是比打我眼前经过的另一个路人更接近。社群性昆虫能有这种检测亲缘关系的优越权力吗？最简单的社群性动物采取的一个聪明的解决办法是，单独建巢，让你的后代成为你的员工——但如何才能让它们确信你是它们的妈妈而不是一条寄生虫，一个伪装成妈妈的闯入者呢？

大约20年前，化学家将社会生物学家邀请到了化学实验室。于是有关社群性昆虫如何识别亲缘关系的机制研究开始有了一些进展。香水和化妆品，什么手段都使上了。昆虫表皮上的碳氢化合物使个体能够辨认伙伴。科学家们将其他蜂巢的香味涂抹在黄蜂身上来察看同伴的反应，结果同伴表现得异常恐惧，就像面对一个入侵者。但是化学物质也并非万无一失。寄生蜂就没有什么气味，这使它们可以悄无声息地入侵宿主的巢。随着时间的推移，寄生蜂获得了宿主巢的气味，于是巢内的工蜂开始为入侵者工作，因为它们无法区分谁是真正的伙伴。其他科学家则将胡蜂的脸上打上眼线和口红，这样也能切实搞乱黄蜂的识别系统。

那不等于说昆虫识别亲缘关系是靠气味和长相，而不是靠与生俱来的平衡亲疏指数的天赋？科学家们能确定的是，社群性昆虫确实可以从非同巢同类中识别出同巢的伙伴来，正如你知道谁是你的朋友和邻居。但是社群性昆虫可能不像我们想象的那么非凡超绝，可以为巢内的每个个体或成虫标上亲缘指数。有人认为，某些高度进化的物种，如某种蚁类，可以做到这一点，但这方面还缺少确定性的实验证据。当然，即使是最聪明的昆虫也会犯错误，但我们不打算过多地讨论这方面内容了。人们发现，蜜蜂、大黄蜂、黄蜂（vespine wasps）和胡蜂（polistine wasps）经常会在“错误的”蜂巢里工作。事实上，我们观察得越多，就越是会发现这些“错误”。让我们帮汉密尔顿一个忙，看看蜂箱外的情形，这样你对这些“错误”会有新的（近亲选择）的解释。

困境3：何时成为社群中的一员

我们的英格兰蜜蜂有理由抱怨英格兰糟糕的天气。毕竟，是天气影响了它对群居还是独居的选择。对它来说，选择群居还是独居的决定必须在开春的时候就做出。它的法则也许是这样的："如果到4月中旬气温还不到10℃，最好是选择单过，因为面临的将是一个寒冷短暂的繁殖季节，没时间孵化两窝。"如果确有这样一个法则，它怎么会经常被打破呢？难道处于英国南部温暖季节的孤独雌蜂总犯错误？

我在巴拿马做了很多这方面的研究，在那里我享受着一种均衡的、不太忙碌的生活。那些生命短暂、难见阳光的英国蜜蜂真该在巴拿马好好过个假期：它可以轻松下来，自由自在地生活。打一开始它就可以自己单过。当它开始产下第一窝卵后，它可以决定是否要把它们踢出去，仍保持单过；或选择过群居生活，把它们全留下来；或终于厌倦了孤独打算与沾亲的邻居共同生活。亚热带蜜蜂与温带纬度蜜蜂不同，她们的生活终其一生都具有可塑性，因此她们可以随季节环境变化采取不同的生活方式。也许只有在温带地区才存在永久性的形态级别演化。

我们需要知道如果蜜蜂作出了错误决定将会发生什么情形。它的灵活性如何，在许多简单的蜂群里，蜜蜂和黄蜂都可以从工蜂变为蜂后。就是说，对雌蜂来说，不论是在自己家的巢里还是在邻居的巢里，作为工蜂只是一种暂时性的选择，它们在等待一个更好的选择。我们的蜜蜂联盟会解体吗？它会被裁减下来去过一种孤独的生活吗？对于具有高可塑性的雌蜂来说，跳槽到附近的一个阳光更加充沛的巢里去生殖是一个可行的选择。

那么在哪个关键环节上社会性昆虫失去这种可塑性了呢？对有些种类来说，是寿命限制了这种品种的可塑性。难道这个原因就一定导致演化上的永久等级制？生态学因素会不会也是一种制约因素呢？现在看来，还没有发现一种单独占用的蜂巢。如果我们能够从基因方面、环境和生态等方面找到控制蜂群等级行为的原因，并在蜜蜂失去可塑性时确实观察到这些作用，那么就能够拨开这团迷雾。

最终水平上的社群性进化

经过对像蜜蜂和蚂蚁这样的高度进化、复杂的社群性昆虫社会的几

十年的观察，我们实际上对这种社群性的起源已经知道很多。关于蜜蜂是如何获得社群性的，它们何时以及为什么会这么做，它们是如何处理汉密尔顿法则所说的亲疏关系的等问题，我们多少知道一些相关的机制。现在重点要解决的是这种社群性的起源问题，并为它确立坚实的基础。

无论喜欢与否，眼下生物学家都急急地奔向各种“组学（-omics）”的前沿：基因组学、蛋白质组学、基因转录组描述（这些学科分别研究基因、蛋白质和各种基因的表达）。好像什么都可以构成“组学”：metabolomics, interactomics，连“Sociomics”——一种研究基因和分子是如何构成社会系统的学问——都出现了（见http：//omics.org）。太好了！基因组巨人已在生物学领域显身，它剥去了自然世界那种古典的优雅，将它还原为赤裸裸的基因组，使那些久藏的奥秘被揭露出来，原来它是那么脆弱。通过窥视蜜蜂的基因组，我们可以找出社群性昆虫的祖先。通过检查在蜜蜂生命周期的不同阶段里“孤独的”与“社会的”基因是如何表达的，我们可能最终会逐渐发现这些祖先到底是锁定在某个孤立世界里隐秘的社会活动家呢，还是很不愿意参与社会活动的隐士。也许它们两者皆不是，也许它们已演化得相当灵活，可以按环境条件采取具体的生活方式。如果我们能够解决这个问题，那么我们确实寻找到了社群性昆虫的祖先，实现了最终突破。

借助于基因组巨人及其在诸如果蝇、线虫以及最近的蜜蜂等模型生物体方面的工作，我们可以通过研究基因来理解昆虫社群性的起源，这使我们的认识提升到一个全新的水平。不仅如此，现在我们可以对过去想都不敢想的一些新问题进行讨论。例如，昆虫在向社群生活过渡时其基因组是如何变化的？是不是某些基因被复制、改组并发生突变？如果是这样，那是哪些基因呢？哪种“孤独”基因可能会丢失？新的“社会”基因能演化出全新的功能吗？在社群性刚出现的时候，基因的可塑性是如何被刚性、可靠性和持久性所取代的？

我希望现在我已经让你相信社群性昆虫是非常丰富和好玩的，我们才刚开始了解它们的起源。如果真要将它们的生活与我们的生活进行类比的话，可以说，经过8000万年的演变，社群性昆虫无疑是聪明和神奇的。如果要自我遴选动物王国的领袖，我们只有谦卑地站在社群性昆虫面前的份儿。

本文开始时我将昆虫社会与人类社会做了平行比较，这表明我们可以从它们那里学到很多东西，劳动群体成员之间的明确分工就是一个明证。从某些方面来看，我们并不是最好的学生，例如：像当今许多富于进取的人类社区里的妇女一样，我力图做到生育（我的“蜂后”的一面）、工作（在科学世界里打食）两不误。人类最好吸取来自社会昆虫的经验（毕竟人家比我们早了几百万年），将劳动分工迅速由性别决定转移到由技能决定上来。当然这完全是另一篇文章了。

第17章
人类的灭绝和进化

◎卡捷琳娜·哈尔瓦蒂

卡捷琳娜·哈尔瓦蒂（Katerina Harvati）

古人类学家，专业方向是尼安德特人的演化和现代人类的起源。她的研究兴趣包括演化理论、形态变异与遗传、环境等因素之间的关系，以及灵长类和人类生活史的演化。她主要在非洲和欧洲包括她的祖国希腊进行实地考察研究。

哈尔瓦蒂曾在哥伦比亚大学、纽约城市大学和美国自然历史博物馆从事研究。在2004年加入马克斯·普朗克演化人类学研究所之前，她是纽约大学人类学系的助理教授。她还是纽约城市大学研究生院人类学副教授。哈尔瓦蒂和她的同事对南非东开普省发现的晚更新世人类头骨的研究曾被《时代》杂志评选为2007年度十大科学发现之一，她们的研究提供了“现代人类是在65 000年前到25 000年前之间离开非洲的首次化石证据”。[54]

这在描写灭绝的文学作品里几乎是陈词滥调：查尔斯·达尔文和那些最突出的进化论思想家没有太多地注意到灭绝其实是进化的重要力量。想到地球上曾存在过的物种绝大多数已不再与我们同在，这不免有点令人惊奇。对灭绝过程进行认真研究只是在过去几十年里的事情——人类现在对地球的影响已变得十分明显，我们正面临着生物多样性的危机。

今天，即使是粗略地瞄一眼2008年世界自然保护联盟列出的濒危物种红色警示名单（http：//www.iucnredlist.org），就能深切地感受到当前物种灭绝的危机以及它对我们自身的生存前景的影响有多么严重！如果我们检查一下化石记录，这种印象会进一步得到强化。在通常随时随

地发生的零星灭绝事件“背景”下，生命演化史上曾发生过数次大规模的灭绝事件。有些事件曾严重到使当时全部生物种群的一半以上遭到毁灭，导致地球的生物群落完全重建。这些事件中最广为人知的（尽管它未必是最严重的）也许要数K/T（白垩纪/第三纪）事件了——大约在6500万年前，恐龙灭绝了，随之而来的是哺乳动物的繁衍和进化。

虽然在我们这个物种的生命周期里没有发生过如此激烈的大规模灭绝事件（也许只是到现在为止），但物种灭绝对于我们的进化却十分重要。这一点从人类化石记录中两个相当明确的灭绝事件就可以看出：一个是约100万年前傍人属（Paranthropus）的灭绝，另一个是3万年前尼安德特人的灭绝。这两次灭绝在人类进化史上都具有重要的标志性意义。

生物种群为什么会灭绝？

这个问题的答案似乎是显然的，几乎不值得考虑，但事实是，我们确实不知道。灭绝的原因相当复杂，并没有得到充分认识。在目前状况下，人类活动的直接或间接影响不仅是一个重要因素，而且其重要性已远远超过以往任何时候。人权机构的一些研究者声称，在最近的化石记录的数起灭绝事例中，最突出的要数约5万年前澳大利亚巨型动物的灭绝和1.1万年前的美洲巨型动物的灭绝。这里除了智人（Homo sapiens）的出现在时间上相距较近之外，引起物种灭绝一定还有其他重要因素。更重要的是，这些因素导致了整个地区的动植物的消失。

关于物种灭绝的原因主要有两种观点。第一种观点（达尔文所青睐）看重生物间的竞争，被灭绝的物种与生存下来的物种之间往往有密切的亲缘关系。由于自然选择产生出优良品种，新物种最常见的是改善了它们祖先的版本。当两个有亲缘关系的物种发生接触时（由于大陆板块重新合拢的缘故，原先彼此隔绝的具有亲缘关系的物种又出现融合），一般总是子代在竞争中胜出，从而导致旧物种的灭绝。这种观点导致了竞争排斥原则的确立，它对于解释如下情形是成立的：在一个地区内，具有相同生态要求的两个物种不可能长期共存，其中一个最终会导致另一个的灭绝。其原因可能是胜者内在地就具有更有效利用资源的优势，或者即便没有竞争优势，但它有更大的初始种群。虽然我们经常谈及竞争就想到侵略，但每当出现两个物种都需要从环境中获取同样的

有限资源（譬如食物和栖息地）的情形时，这一法则就会起作用。入侵物种的行为只是极端的例子，这类入侵可以从化石记录中知晓——例如，当大约350万年前南北美洲大陆在巴拿马地峡合并时，就发生过北美生物群向南美迁移的事例。这次迁移被称为美洲生物大交流，这一过程导致大部分南美本地的动物灭绝，取而代之的是从北方侵入的物种。人类作为外来物种的侵入，也同样引起本地动物群的灭绝。

物种灭绝的另一种常见的解释是环境变化，通常是指气候变化。这种观点认为，物种灭绝是一种带有偶然性的事件，一个物种被灭绝不是因为这个物种的生存能力较差，而是因为运气较差。从某种意义上说，它出现在了错误的地点和错误的时间段。这种解释将原因归咎于火山活动增加、海平面变化、大陆板块的位置移动以及地球轨道参数的变化等，所有这些都可能引发气候变化或栖息地的丧失。这样的地质事件在地球历史上层出不穷，因此可以很好地解释背景灭绝的发生——但由于这些事件发生的频率很低，而且往往具有相当长时间的渐进性质，因此它们不能完全令人满意地解释物种大规模灭绝的原因。有人认为，即使是单个物种濒临灭绝，那种突然出现的、灾难性的变化也是必不可少的条件。因为环境变化必定极为罕见，以至于在物种进化史上是不可预测的，而这类事件又是如此突然，使得物种没有时间通过自然选择行为来避免灭绝。这种环境灾难的极端表现当属极为罕见的小行星或彗星撞击地球引发的自然灾害。人们强烈认为，白垩纪第三纪期间的大规模物种灭绝就是由此引起的，化石记录下的其他类似的大规模灭绝也可能与此有关，虽然这种说明可能会受到由于沉积物保存的化石质量欠佳的影响。

并非所有物种都同样容易地被灭绝，无论起因是竞争还是环境变化。生物种群虽然不具有特定的性状来保护其免遭大灭绝，但有些特征还是能够在不太严峻的环境下提供一定程度的保护。大范围的地理分布就是一种有利因素，因为数量庞大，高度分散，使物种进化出具有忍耐和广泛利用环境的能力。而身体过于庞大、妊娠期和成熟期过长则加剧了物种的脆弱性。孕期和成熟期的长短往往是相互关联的，主要由狭窄的适应性条件决定。也许不利因素里最重要的是物种在数量上过小。岛生物种特别容易迅速灭绝，正是这个原因。

不是所有的人类化石都是我们的祖先

这一观点——人作为一个物种可能会灭绝——还没有被广泛接受，并且仍然面临着来自古人类学界的重重阻力。与这一观点结伴产生的一种认识是：过去可能同时存活着两个（或更多个）人类类群。这种情形与我们的经验非常不同，我们的经验是建立在这样一种当前认识基础上的：智人是一个独一无二的、广泛分布并且体貌特征随地理环境不同而不同的物种。在20世纪50年代，进化论思想家曾以目前条件为准则提出了单种假说。这一假说是将竞争排斥法则运用于人类（化石记录下的人类种群）发展而来的。它假定，人类祖先一经发展到使用工具和有了文化，这两者就成了他们适应生存的独门绝技，使他们能够广泛利用环境来扩大生存空间。由于文化是人类特有的适应手段，因此根据竞争排斥法则，同一地域同一时间上不可能并存两个有文化的人类群。此外，由于文化具有提高人耐受环境压力、适应地理差异的作用，因此就物种生成而言，人类世系的长期隔绝几乎注定是不可能的。

这种假说几乎可以毋庸置疑地运用于说明离现在最近的更新世的化石记录，也可以延伸到说明较早时期的南方古猿。它消除了人类世系的产生和灭绝的问题。既然可以假定一定时间段内只存在一种人类，你可以进一步假设它不可能已经灭绝（至少在终端，即种系的意义上是如此——也就是说，没有演变为新的物种），因为我们在这里！这种处理问题的方法将化石人类置于一种特殊的情境下，从根本上移除了它在自然世界里的位置，直截了当地把它置于我们所谓的现代人类的保护伞下。

自从单种假说提出以来，古人类学已经走过了漫长的道路。这一领域的进展日益指向人类进化的丛林模式，就是说，多分支人科物种或许符合我们的传承准则。现在很清楚，人类（无论是化石人类还是当前的人类）并不能免除生物力量的制约，灭绝只是早晚的事。

傍人属案例

按照人类化石的记录，傍人属（有时也被称为南方古猿粗壮种）大约出现在270万年前的东非和南非。虽然它像我们一样有双足，但其大脑与猿类没有两样。其头部形态与它具有功能强大的咀嚼器官有关：外倾的颧骨，几乎凹下去的面部，颅顶的矢状脊，大的颊齿和下巴。这种极端的形态显示出，傍人属的食谱非常单一，因此在解剖学上不属于更古老的南方古猿群类的变种，这对于人类进化早期的单种假说是个决定

性的打击。

傍人属的颅骨形态差异是如此之大，为此将它们单列为一个独立的属，它包括3个种，所有这3种都与晚期的人科物种相差甚远，我们很难将后者看作是从前者演化来的。虽然化石记录带有一定的不确定性且断代上并非总是正确无误，但从傍人属的最晚近的标本我们大致可以断定，它们出现在距今约100万年前。傍人属的灭绝已无可争议，问题是这种灭绝是什么原因造成的？是什么力量造成了这个属的早期分化？

傍人属的起源大致与我们自己这个属的早期代表——人属——的出现处于同一时期，在这一时期，非洲已出现早期石器，气候变化使得土地变得越来越干旱。人们普遍认为，这一气候变化使得南方古猿演化出两个支脉：傍人属和人属。直到最近，人们一致认为，傍人属通过进化出强健的咀嚼器官适应了新环境，而人属则遵循不同的路线，通过发展石器工具技术（因此文化在这里起着作用）扩展了生存的适应性。因为只有这两个属之一的人属掌握了石器工具技术，这两个人科动物共存才并不违反排斥法则。这个方案预设了人属是早期工具的唯一制造者，但这一假设受到了质疑。傍人属也拥有制造工具所必需的灵巧性，并被认为制造过用白蚁捕鱼活动所需的骨制工具。

也许正因为狭窄的生态适应性——就食性而言——才容许傍人属与食物结构宽泛得多的早期人属共存。由于食谱单一的物种比广谱食性物种更容易灭绝，加上100万年前冰期的频繁循环，因此食谱单一性可以解释傍人属的消亡。然而，有最新证据暗示，尽管傍人属有着令人印象深刻的咀嚼器官，但无论是在食物结构上还是在栖息地偏好方面，它绝不比早期人属更单一化。事实上，在某些抗打击能力方面，傍人属的得分要更高些：它是迄今为止化石记录中这一时期最丰富的物种群类，这表明当时傍人属具有广大的种群数量。其种属之一的南方古猿鲍氏种（P.boisei）是最长寿的，因此也是进化上最成功的人科物种之一，其存在的时间跨度超过100万年。

导致傍人属灭绝的原因尚不清楚。这些原因可能包括环境的剧烈变化或与另一类群——很可能是智人——的竞争。但可以肯定的是，大约在100万年前，傍人属在化石记录上消失了，只留下早期智人作为我们传承的唯一代表。

亲缘种

与傍人属不同，尼安德特人与我们十分相近，属于体型大、脑容量大的人科，并具有复杂而记录完好的文化行为。他们生活在距今大约3万年前的冰河时代的欧洲。像我们一样，他们是猎人和肉食动物，用火，并埋葬死者。但是他们也有明显区别：在用以比较灵长类动物的框架内，尼安德特人的头盖骨与现代人类的头盖骨之间的解剖学差异要比两种黑猩猩之间的差异大。现在普遍认为，尼安德特人是一个独特的物种。大约在50万年前，他们从我们的共同祖先那里分叉出去，并从非洲迁徙到欧洲，变得与世隔绝。地理上的隔绝可能是造成两个人类物种同时进化的原因。虽然50万年是一段相当漫长的时间，但从进化的角度看则还很短暂。尼安德特人是我们最亲近的亲戚，它与我们的亲缘关系要比与黑猩猩的关系亲近得多。他们是我们的姊妹种。

大约在3万年前，尼安德特人从化石记录中消失了，没有留下后代，现代人类的基因库中也找不到它的痕迹。现代人（智人）大约在4万年前到达了欧洲，几千年后尼安德特人就灭绝了。因此，显而易见的问题是：尼安德特人的消失能像很多其他物种的灭绝那样归咎于环境变化吗，还是智人的到来（带来的某种暴力）使然？

这里，突发性的、灾难性的气候变化的情形很容易排除。在尼安德特人消失前后的这段时间里，气候变化从没有剧烈到可以造成这一物种灭绝的程度。[55] 尽管如此，气候的剧烈不稳定性有可能引起尼安德特人的人口数量逐渐减少。但是仅靠气候的压力似乎不足以给出令人满意的解释，毕竟，尼安德特人在欧洲是进化了的，并成功抵御过以前的剧烈的气候变化。

但气候压力可能是导致尼安德特人与入侵的近亲物种发生激烈竞争的原因。我们能找到这种竞争的记录吗？没有任何证据表明智人与尼安德特人之间发生过侵略性的相互作用——但竞争要比侵略更经常地发生。现代人捕猎的对象与尼安德特人的一样，都是大型草食动物。如果这种食物资源有太多的重叠，竞争排斥法则可能就会起作用。经过足够长的时间，两者之一就可能驱使对方走向灭绝，即使这种情形出现在大如西方欧亚大陆板块亦不可幸免。

即便如此，尼安德特人也曾取得过一些对我们祖先的进化优势。他

们的成熟期可能会稍快于现代人，因而繁殖周期较短。他们还具有安居的优势：经过千百年欧洲冰河时代的进化，他们在体力上能比新移民更好地适应当地的环境条件。但尼安德特人也有弱项。他们的食谱似乎过于单一，几乎只以大型食草类动物为食。重要的是，尽管生殖周期较短，但他们的人口规模似乎一直不大，他们的受伤比例和死亡率一直居高不下。

现代人能够延续下来可能是因为他们的食物结构更灵活，他们还具有技术优势。虽然早期智人像尼安德特人一样，经常以大型哺乳动物为食，但他们也捕猎不易捕捉的小型哺乳动物、鸟类和鱼类。现代人食谱的多样化可能能够解释为什么这两个物种能够共存于同一地区长达几万年。但我们祖先的最大优势恐怕还是人口。今天我们依然独享着这一人口特征：我们的妊娠期比类人猿和早期人科动物要长，因此我们有非常长的寿命。与猿类不同，我们在两次受孕期之间的间隙时间较短：在发达国家，一个母亲可以在不到一年的时间里生下两名婴儿。即便在现今以狩猎为主的社会里，这也只要花大约3年时间。而黑猩猩母亲则需要每隔6年左右时间才能再次生产。人的这一特征与我们人类的寿命较长有关：通过让老年人，特别是妇女，来照看幼儿，使育龄妇女可以生产更多的孩子。[56] 虽然我们不知道早期现代人的妊娠间隔时间是否比尼安德特人的更短，但我们从化石记录中确实看到，早期智人中老年人的比例要比尼安德特人高得多，这意味着早期智人的人口特征与我们现代人相似。这样的人口特征有助于人口的增加，考古记录也证实了这一点。人口的这一特征还有助于我们的祖先在过去的6万年里散布到全世界各个地方。仅此一点就足以压垮小而稳定的尼安德特人的人口，根本无需一次性打击。

九九归一

现代人在欧洲取代了尼安德特人，可能也取代了走出非洲的智人的所有其他后裔。甚至在到达欧洲之前，我们的祖先已到达澳大利亚。几千年后，他们在美洲建立起自己的家园。这些大陆以前都不曾有过人科动物。我们的广谱食性，我们在各种恶劣环境下的生存本领，我们特有的人口特征，使得我们在短短6万年——按地质年代的尺度衡量，甚至短于一次心跳——的时间里产生了大量人口，并将这一种群扩散到世界的各个角落。现在我们已无处可去。我们是我们这一支脉的唯一幸存

者。我们的数量超过了狮子，我们的活动已透支地球的生态系统，很可能还改变了地球的气候。其后果是使得越来越多的生物感到害怕，但它们对我们来说意味着什么呢？我们造成的新的生存条件能够维系我们的存在吗？如果我们不小心处理，我们是不是就将面临与过去的竞争对手同样的命运？有一点似乎很清楚：我们必须再一次调整我们的行为，以应付规模空前的气候变化和接近大规模灭绝的生物多样性危机的挑战。好在我们还有时间来适应。

第18章
为什么专业化没有导致科学四分五裂？

◎加文·施密特

加文·施密特（Gavin Schmidt）

美国航天局纽约戈达德空间研究所的气候学家。他的工作是模拟过去、现在和未来的气候。他于1989年在牛津大学获数学学士学位；1994年获伦敦大学学院应用数学博士学位。随后，他去蒙特利尔吉尔大学做博士后直到1996年。这之后，他获得国家海洋和大气管理局气候和全球变化博士后奖学金，来到戈达德空间研究所。

施密特曾是入选《科学美国人》杂志2004年度50位最杰出的研究者之一。他是2007年度诺贝尔和平奖联合国政府间气候变化专家小组获奖报告的作者之一。他还是RealClimate.org网站的创始人之一和特约编辑。这个网站提供了大量有关环境和气候科学问题的背景知识，这些背景知识往往为大众媒体报道所忽视。

科学家是这样一群人：他们对那些越来越没人关注的知识知道得越来越多，以至于他们通晓一切。

——约翰·齐曼

还原论大行其道。两千多年来的科学进展表明，将问题分解成越来越小的组成部分以看清复杂现象的本质是了解自然世界的一种很好的方法。然而，随着科学文献数量的成倍增加，人们也更加关注这样一个问题，那就是日益专业化已使得不同领域的科学家很难进行交流，更不用说合作了。

尽管这种情况很可怕，但跨学科科学已经成为并将继续成为一支重要的认识世界的力量。那种将还原论发挥到极致的力量正受到来自跨学

科研究的有力反击。这里，我将通过展现新的专业是如何创建以及它们是如何将信息反馈到一般科学领域的过程来说明科学因素和社会因素是如何在其中起作用的。

到18世纪末，一个人想跟上所有新科学的进展已经不可能了。有人认为，物理学家、医生兼埃及学家托马斯·扬（Thomas Young，1773—1829年）是最后一位“通才”，也有人认为这份荣誉应属于考古学家、数学家、生物学家、物理学家、火山地质学家兼埃及学家阿塔纳斯·珂雪（Athanasius Kircher，1602—1680年）。其实谁是最后一位全能型博学者并不重要——或者说，如果这样的人确实存在过的话。但是，随着19世纪初科学期刊的兴起，科学家开始抱怨研究的分岔并对跨学科研究机会的日渐减少表示出强烈的不满。人们指责越来越专业化已成为几乎所有科学领域的通病。这一点从出席1900年美国科学促进协会波士顿会议的人数的减少到1959年斯诺在他著名的瑞德讲座《两种文化》[57] 中所讨论的人文与科学的隔阂问题上，都有反映。

虽然科学并没有分裂为相互间不和谐的杂拌儿，但确实存在将科学景观割裂开来的各种力量。这些力量有些是自然的，不可避免的，而且值得赞赏；有些则是体制上的，可以避免的，并应该受到谴责，而且事实上也存在着打破任何僵化分类的强烈冲动。下面我用离心力这个词来形容那种使科学分支越来越分化的力量，用向心力一词来形容反抗这种分化的力量。

我们可以将科学活动看作是一个在离心力和向心力共同作用下推动其前行的动态系统。我的这一观察来自我自己在地球科学方面的经验——具体而言，是气候和气候变化研究方面的经验。对这一广泛领域做一简要概述可能有助于后面的讨论。

地球的气候系统由地球大气、海洋、陆地和冰川组成：它们的动力学、化学和结构以无数种方式混合构成了“气候”，即日常意义上的“平均天气”。科学家采用各种不同方式来研究这个系统：直接观察热带太平洋的变化（譬如厄尔尼诺事件），分析卫星测得的海平面变化的数据，测量南极冰盖下80万年前形成的空洞的气体浓度，并为此开发出各种复杂的大气环流计算模型，不一而足。

所有这些努力要回答的基本问题是相似的：气候为什么会是这样？

过去它是怎样变化的？现在有什么变化？将来又会如何变化？为什么会这么变化？等等。但研究的时间尺度、采用的方法以及得出的答案则是千差万别。由于所有气候科学家研究的是同一件事情，因此尽管搞模型的、搞气象的和搞冰核的专家提出问题时采用的概念不同，但得到的答案应该类似。因此，考察跨学科科学在气候变化研究领域如何进行是观察向心力和离心力如何实际影响科学的一个好方法。

为什么各专业领域会离得越来越远

就像胆固醇，离心力也有好有坏。好的离心力源自无懈可击的还原论动机：需要尽可能减少造成臭氧层耗竭的化学反应，需要明确阐明海冰热力学的具体细节，需要改进大气辐射传输模型。这其中的每一项工作都需要全力投入，需要具体的各不相同的实验室技术、实地观察技术或复杂的数学知识。对这些研究者来说，没有任何特别的理由非要将这些问题放在一起讨论。实际上，对他们每天面对的问题、任务和解决方案感兴趣的主要是他们的同事。坏的离心力则来自我们人类的弱点：专业意识、惰性、官僚主义以及可以理解的职业安全的愿望。

新的专业领域常常孕育出新的工具或技术，这些新技术已经产生出有趣的结果。其他科学家会接着进行深入研究。如果这种工具或技术显示出强大的生命力，那么用不了多久学界就会组织会议来专门研究这些成果，以此为目标的攻读博士学位的项目也会很快设立，教授职位和专业期刊也将随之而来。随着这种发展变化的出现，人们还会创造出新的术语，并将新领域中的这种成功变成样板。这样的例子不胜枚举：PCR（聚合酶链式反应）带来了DNA指纹技术、基因测序技术和人类基因组计划；计算机硬件的发展带来了预测天气的数学模型；卫星应用开辟了遥感领域；欧洲核子研究中心（CERN）的超级对撞机则为高能物理研究带来了全新面貌。这些是好的方面。

最近的一个例子是在地球科学基础上诞生出的古海洋学分支——一门研究海洋过去的变化的科学，这门学科的研究内容涵盖了关于过去一亿年间海洋环流、海洋化学以及生物学等方面的变化。导致这一研究领域出现创新的是深海钻探技术的发展，它使科学家几乎能够从海底任何地方提取古海洋淤泥样品。使这一技术成为新学科研究手段的科学突破是20世纪60年代的一项发现：地球化学示踪设备在海底泥面薄壳采得的

时间序列显示，冰期呈现周期性盈亏现象——而且这种现象已持续了数百万年。人们从每一海盆测得的关联数据的分析中得出结论（当然还借助了其他一些证据）：冰期的周期性起因于地球轨道的可预见的摆动。

正如所料，新的技术词汇出现了——轨道调整、有孔虫传递函数、烯酮测温、方解石溶解指数、钍校正沉积速率等。更微妙的是，新学科借鉴了其他学科的许多词汇并赋予其新的含义："高分辨率"在这里指的是沉积速率大到足以分辨出十年或一个世纪的气候变化，这个定义与物理海洋学、天文学或微生物学中使用的定义有很大的不同；深海的"通风"也与空气流动无甚关系，而是指深层海水与表层海水混合的各种行为。

对新技术的使用者来说，从现有学科领域拿来各种术语和概念用以说明研究了几十年的系统是非常有意义的。但由于用新技术看到的东西不一定是这个概念原有意义下的东西，因此同一个概念在应用到略微不同的现象上时其原有的意义已告终止。这让我想起印度的盲人摸象的寓言——不同的人摸到不同部位后就认为大象像绳子或像墙壁或像长矛等。虽然这个寓言反过来形容也许更合适：不能仅仅因为摸到了绳子，于是就认为这是一头大象。

正是在这个地方人们开始看到了使得科学分支相互隔离的不科学的因素。一个很好的例子是横跨热带太平洋的海水温度梯度的变化。按现代的观察，这些变化都与厄尔尼诺事件有关，所谓厄尔尼诺是指西太平洋暖流向秘鲁沿岸的动态漂移。对历史上海水温度梯度变化的观察也都伴有类似的活动，但这种"古厄尔尼诺现象"却未必能得到现代海洋学家的认同。套用温斯顿·丘吉尔的话就是，专家会因为日常语言而变得疏离。

另一个社会因素是，随着科学成果的积累，会出现一种特定的亚文化和一整套如何在新领域获得成功的模式。年长的上一代人总想按自己已有的模式来录用和提拔接班人，但年轻的一代则想超越前辈。在这两种情形下，都会遵循一种以争取业内同行一致好评的工作模式。在古海洋学领域，这种模式是由40年前开始进入这一领域的一批数据生成地质化学家建立起来的。他们的声誉是基于他们扎实的数据分析能力——他们具有将海洋上不同地点采得的数据联系起来的独门（或近乎独门的）功夫。海洋钻探取得的岩芯必须能提供特定地点的时间序列，问题是难

在对这些岩样进行细节比较，因为采得岩样的深度与年代对应上存在着很大的不确定性。这样，为特定岩芯发展出一套“年代判断模型”往往就成为适当归置新数据工作中最具挑战性的部分。因此在这一领域，发表的论文绝大部分讨论的是特定地点采得的岩样的明显变化，而不是更难评估的某个时间段上岩样的地域变化规律。对于有抱负的年轻人来说，古海洋学的研究路径是明确的：寻求海洋钻井公司的支持，购置高精度测量设备。这些设备通常是一些质谱仪，耗资大约100万美元，而且往往成为聘用新教师谈判的一部分。这些设备也带来了责任：为技术人员提供资金支持，拿到有趣的岩样，以证明本实验室的这些钱没白花。

因此，年轻的研究人员被锁定在该领域主流的工作模式上。他想在职业生涯里从事不同类型的科学工作是困难的。具体到古海洋学领域，当前研究的热点是海洋芯部岩样的垂直比较，重点在如何收集和分析样本，可你偏对岩样的横向比较（即不同地点同一深度上的岩样变化——译注）感兴趣，这就有问题，因为横向比较所需的信息分布得极其广泛。通常人们不鼓励这种跨学科的努力，因为对于气候变化的地理模式这种“横向”问题来说，更广泛的气候学领域的研究更有效。

总之，一个学科分支一旦建立起来后，其必要的和特有的专业性便会对跨学科研究形成一种障碍：各种专业词汇使得该学科外的大多数人不知所云，从其他学科借鉴来并赋予了新内涵的术语会在对外交流时引起微妙的误解。专业期刊、各种会议、新成立的院系、对新设备的投资和宣传会接踵而来，这些举措加深了对新方法的应用。所有这一切会形成一种孤立的亚文化，其交流将越来越远离更广泛的科学领域。从业人员做的都谈不上错，但是，正像经济学里的“公用品悲剧”那样，单个理性人的决定的集体效应会导致一种学科整体所不希望的结果。这一结果反过来又可能导致外部世界不能正确鉴赏其成功经验，并导致其研究经费的削减；从内部来说，则是感到一种缺乏“认同”的挫折感。

迈向综合的途径

为了了解科学整合工作的向心力作用，我们不妨从某个相关领域的视角来观察一个学科。该领域的研究人员是如何发现本领域的新成果的呢？他们如何看待这些成果呢？促使他们从事这一工作的更深层次的动机是什么？

新的科学发现的传播主要有三种途径：科普活动、“全权大使”和技术文献的出版。对于第一条途径，说来你可能会吃惊，譬如许多气候科学家都读过詹姆斯·格雷克的畅销书《混沌》而不是爱德华·洛伦兹关于这个问题的原始文献。至于我自己在科普方面的尝试，很多同事对我说，他们发现我关于一些重要细节问题的非专业描述非常有用，这令我感到很诧异。

第二个途径是指具有某一个学科专业背景的科学家愿意花大量时间向更广阔领域的科学家介绍本学科的成果。这些科学家往往是些愿意跨学科交流的人，他们强烈提倡各领域间的合作，因为他们能够在不同领域间进行有效的沟通。但这样的人现在是越来越少了。

至于专业文献，科学家和普通市民没两样，通常认为不可能读懂（更不用说判断其价值了）不同领域的最新论文。他们关注的文献大致可分为三大类：首先是那些出于内在兴趣而关注的真正的大问题研究——例如，路易斯和沃尔特·阿尔瓦雷斯于1980年写的关于白垩纪晚期小行星活动对白垩纪结束的影响的论文，或凯特·斯彭斯于2000年发表的关于用地球轨道进动和经验法则来判断古埃及正北因

而与金字塔建造时间无关的研究等。然而，尽管这些论文具有标志性意义，但却不一定是多方合作的结果。

受关注的第二类文献是原创性的具有更广泛意义的专业论文。例如，那些报告马萨诸塞州以往飓风登陆的文章，或关于6000年前厄尔尼诺事件的文章，或人在上个冰期结束时大型哺乳动物灭绝过程中的作用的文章等。这些研究可能与现今的某一次飓风，或某一次厄尔尼诺现象，或某种生物的灭绝有一定关联，但直接关系通常很小。这些研究的定域性从更大范围来看往往不具有代表性，时间段或时间分辨也可能过于狭窄，取得数据的环境条件可能过于复杂或不够清楚。有时，用语上存在的如前述的细微差别还可能导致沟通上有问题。这类研究通常只是偶尔带来成功的学科间互动——例如，当我们用更标准的历史信息对古记录进行整理时，发现它们确实与其他数据存在衔接。

第三类也是最重要的一类文献是综合性文献——将许多个别研究得出的精髓进行归纳和汇总的研究。（毫不奇怪，最好的大使往往也是最好的综合者。）这些综合研究属于界间研究（metastudies），我们在流

行病学研究中，在对特定的以往的气候数据进行“横向”重构以便将许多离散的记录关联起来的研究中，都可以看到这类文献。通常，这种综合涉及各种数值模型，譬如气象预报就需要用模型将不同类型的天气数据统一起来，以便得出对大气层状态的协调一致的估计。大范围天气变动模型的发展本身就是综合研究的一个重要例证，它汇集了计算科学、大气科学、海洋学和冰雪圈研究专家的各种成果。

好的综合性论文往往为大量的跨学科研究提供了一个方便的跳板。这对于古气候合成研究可谓千真万确。例如，通过绘制上一个冰期高峰期的温度和冰盖的分布图，我们可以将从格陵兰岛和南极洲的冰核记录连接起来，并重构最近几个世纪的气候演变模式。这些文献带动了更大范围的建模工作，以期理解这些模式，它们鼓励人们去探索互补的研究途径。这些文献通常可以在更广阔的群体中显著提高一个学科领域的形象和重要性，这对于获得更多的进一步研究所需的资源以及人们对该领域工作重要性的评价都非常关键。

对科学研究的向心力的探索大致相当于问这种综合为什么会发生和它是怎样发生的。当众多零星的数据或物理过程积累到足够多，使人们很难跟踪的时候，就会出现对综合的需求。大致上说，综合研究关注的是厘清头绪。不同的科学家对这种综合分析的态度差异极大（毕竟，有些东西在一个人看来是噪声，但在另一个人看来则是信号），这种差异在选择从事综合研究项目的人当中也有反映。

综合研究既可以从底层（即领域内）做起，也可以从高端（即在局外人或机构目标的驱使下）至上而下地进行。对最近一期冰期气候变化的估计研究属于自下而上型，由从事古气候科学家团体提出。相反，在海洋学领域，人们经常引用的《世界海洋温度和盐度图集》则是由国家海洋学数据中心的科学家将数以百万计的单个数据点编制起来完成的，而不是由哪一位物理海洋学家自己收集数据就可以完成的。事实上，外部需求是很常见的驱动力。它可以来自建模者的需要，建模需要有综合数据用来与仿真结果或来自其他机构（譬如政府间气候变化专门委员会或国家科研部门）的评估结果进行比较。

尽管这种综合研究具有广泛的影响，对由此产生的学科更是影响深远，但来自领域内的反应则几乎全是批评性的，有时甚至是敌视的。人们常常断言（倒也不是毫无道理），综合研究结果让人根本无法接受，

因为其中有太多的未知因素和太多的混淆视听的因素。还有人抱怨说，搞综合研究的人都是些自己不花力气尽享用别人辛勤采集数据成果的人。而对于综合成果的外界要求，人们则往往抱怨说很多细节被忽略了，这将特别是在公众中导致误解。

有一点可以理解，尽管综合性成果往往难产，但这些成果都经过严格审核，主要是在良好的科学效果方面。例如，建模者第一次给出的冰期海洋冷却的热带范围估计就与陆基温度估计不一致，由此导致了长达数十年的对所采用方法的重新评估，并最终产生新的自下而上的综合。

具有讽刺意味的是，一些获得领域外认可的工作往往在领域内遭到抵制，对综合研究的进一步支持往往很难成为主流。正是在这里向心力面临着最明显的离心力的抵抗，这既可能是好事也可能不是。这值得我们花时间来综合吗？应当说，尽管没有它，我们一样可以在某个领域取得令人印象深刻的进展，新见解一样可以继续涌现出来，但外部评价则可能滞后，由此将引起越来越严重的来自内部的挫折，这种挫折可以成为向心力，促使科学家们合作，以分享他们的发现。而分享成果的主要机制就是综合，这种综合越容易做到，就会形成越多的跨学科科学。

这一过程可以加快或得到鼓励吗？我认为可以。当人们认清了阻碍综合和沟通的离心力后，就会想出专门的办法来对付它们。例如，资助者可以拨出专款用于综合研究，使这些项目不与技术研究争经费；只要心里想着综合目标，数据的获取就会变得较容易；还有就是应强调外部对本想研究工作的兴趣能给本学科带来的益处。

从根本上讲，跨学科科学的驱动力来自于这样一种认识——我们希望对所观察到的现象给予解释，我们寻求的答案不依赖于个别科学领域或具体工具和方法。这些工具和方法都是人构造出来的，它们未必就与自然力相匹配。

致谢

我要感谢Vintage Book出版公司的Marty Asher和Jeff Alexander，他们愿意帮忙出版这方面的书籍。我要感谢我的父亲兼经纪人约翰·布鲁克曼，是他启发我有了出版这么一本书的想法。我还要感谢Sara Lippincott，她的出色编辑使本书增色不少。

[1] S.Solomon et al.，eds.，“Summary for Policymakers，”in Climate Change 2007：The Physical Science Basis Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change（New York：Cambridge University Press，2007），13.

[2] Ibid.“Technical Summary，”70.

[3] Solomon et al.，fig.10：18.

[4] F.Hill and C.Gaddy, TheSiberianCurse（Washington, DC：Brookings Institution Press，2003），4—149.

[5] Hill and Gaddy，86.

[6] 见K.S.Coates and W R.Morrison，*The Alaska Highway in World War* Ⅱ：*The U.S.Army of Occupation in Canada's Northwest* （Toronto：University of Toronto Press，1992）.

[7] A.W.Toga et al.，Mapping Brain Maturation.*TrendsinNeuroscience* .29，no.3（2006）：148—159.

[8] C.D.Frith and U.Frith, Social Cognition in Humans.*CurrentBiology* 17，no.16（2007）：724—732.

[9] J.Barresi and C.Moore, Intentional Relations and Social Understanding.*BehaviouralandBrain Sciences* 19（1996）：107—154.

[10] A.T.Wang et al.，Developmental Changes in the Neural Basis of Interpreting Communicative Intent.Social Cognitive and Affective Neuroscience 1（2006）：107—121.

[11] S.J.Blakemore et al.，Adolescent Development of the Neural Circuitry for Thinking About Intentions.Social Cognitive and Affective Neuroscience 2 no.2（2007）：130—139.

[12] 《引爆流行》，钱清、覃爱冬译，中信出版社，2002年版。——译注

[13] M.D.Lieberman，“Social Cognitive Neuroscience：A Review of Core Processes，”Annual ReviewofPsychology 58（2007）：259—289；W.M.Kelley et al.，“Finding the Self?An Event-Related fMRI Study，”*Journal of Cognitive Neuroscience* 14（2002）：785—794.

[14] L.Q.Uddin et al.，“Self-Face Recognition Activates a Frontoparietal‘Mirror’Network in the Right Hemisphere：An Event-Related fMRI Study，”*Neuroimage* 15（2005）：926—935.

[15] O.Blankeetal.，“Stimulating Illusory Own-Body Perceptions：The Part of the Brain That Can Induce Out-of-Body Experience Has Been Located，”Nature 419（2002）：469—470；V Ganesan et al.，“Schneiderian First-Rank Symptoms and Right Parietal Hyperactivation：A Replication Using

fMRI，”*American Journal of Psychiatry* 162（2005）：1545.

[16] M.lacoboni et al.，“Cortical Mechanisms of Human Imitation，”Science 286（1999）：2526—2528；G.Rizzolatti and L.Craighero，“The Mirror-Neuron Systern，”*Annual Review of Neuroscience* 27（2004）：169—192.

[17] K.A.McKiernan et al.，“A Parametric Manipulation of Factors Affecting Task-Induced Deactivation in Functional Neuroimaging，”*Journal of Cognitive Neuroscience* 15（2003）：394—408.

[18] 例如，见J.Gelernter et al.，“Serotonin Transporter Protein（SLC6A4）Allele and Haplotype Frequencies and Linkage Disequilibria in African-and European-American and Japanese Populations and in Alcohol-Dependent Subjects，”*Human Genetics* 101（1997）：243—246.

[19] J.Kaufman et al.，“Social Supports and Serotonin Transporter Gene Moderate Depression in Maltreated Children，”*Proceedings of the National Academy of Sciences* 101（2004）：1716—1721.

[20] S.E.Taylor et al.，“Early Family Environment, Current Adversity, the Sero-tonin Transporter Promoter Polymorphism, and Depressive Symptomatology，”*BiologicalPsychiatry* 60，no.7（2006）：671—676.

[21] J.D.Greene et al.，An fMRI Investigation of Emotional Engagement in Moral Judgment.*Science* 293，no.5537（2001）：2105—2108.

[22] J.D.Greene et al.，The Neural Bases of Cognitive Conflict and Control in Moral Judgment.*Neuron* 44，no.2（2004）：389—400.

[23] E.Ciaramelli et al.，Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex.*Social Cognitive and Affective Neuroscience.* 2，no.2（2007）：84—92；M.Koenigs et al.，Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgments.Nature 446（7138）：908—911（2007）；M.F.Mendez et al.，An Investigation of Moral Judgment in Frontotemporal Dementia.*CognitiveandBehavioralNeurology*. 18，no.4（2005）：193—197.

[24] J.Haidt.The Emotional Dog and Its Rational Tail：A Social Intuitionist Approach to Moral Judgment.*Psycholigcal Review* .108（2001）：814—834.

[25] P.Singer, Famine, Affluence and Morality.*Philosophy and Public Affairs* 1（1972）：229—243.

[26] S.C.Levinson and D.P.Wilkins, eds.，*GrammarsofSpace*：*ExplorationsinCognitiveDiversity*（New York：Cambridge University Press，2006）.

[27] Levinson，*Space in Language and Cognition*：*Explorations in Cognitive Diversity* （New York：Cambridge University Press，2003）.

[28] B.Tversky et al.，“Cross-Cultural and Developmental Trends in Graphic Productions，”Cognitive Psychology 23（1991）：515—557；O.Fuhrman and L.Boroditsky，“Mental Time-Lines Follow Writing Direction：Comparing English and Hebrew Speakers，”*Proceedings of the 29th Annual Conference of the Cognitive Science Society*（2007）：1007—1010.

[29] L.Boroditsky，“Do English and Mandarin Speakers Think Differently About Time?”*Proceedings of the 48 th Annual Meeting of the Psychonomic Society*（2007）：34.

[30] D.Casasanto et al.，“How Deep Are Effects of Language on Thought?Time Estimation in Speakers of English, Indonesian, Greek, and Spanish，”*Proceedings of the 26 th Annual Conference of the Cognitive Science Society*（2004）：575—580.

[31] Ibid.，“How Deep Are Effects of Language on Thought?Time Estimation in Speakers of English and Greek”（in review）；L.Boroditsky，“Does Language Shape Thought?English and Mandarin Speakers'Conceptions of Time，”*Cognitive Psychology* 43，no.1（2001）：1—22.

[32] L.Boroditsky et al.，“Sex, Syntax, and Semantics，”in D.Gentner and S.Goldin-Meadow, eds.，*Language in Mind*：*Advances in the Study of Language and Cognition*（Cambridge, MA：MIT Press，2003），61—79.

[33] L.Boroditsky，“Linguistic Relativity，”in L.Nadel, ed.，*EncyclopediaofCognitiveScience*（London：MacMillan，2003），917—921；B.W.Pelham et al.，“Why Susie Sells Seashells by the Seashore：Implicit Egotism and Major Life Decisions，”*Journal of Personality and Social Psychology* 82，no.4（2002）：469—486；A.Tversky and D.Kahneman，“The Framing of Decisions and the Psychology of Choice，”Science211（1981）：453—458；P.Pica et al.，“Exact and Approximate Arithmetic in an Amazonian Indigene Group，”*Science* 306（2004）：499—503；J.G.de Villiers and P.A.de Villiers，“Linguistic Determinism and False Belief，”in P.Mitchell and K.Riggs, eds.，*Children's Reasoning and the Mind*（Hove, UK：Psychology Press, in press）；J.A.Lucy and S.Gaskins，“Interaction of Language Type and Referent Type in the Development of Nonverbal Classification Preferences，”in Gentner and Goldin-Meadow，465—492；L.F.Barrett et al.，“Language as a Context for Emotion Perception，”*TrendsinCognitiveSciences* 11（2007）：327—332.

[34] H.P.Davis and L.R.Squire，“Protein Synthesis and Memory：A Review，”*Psychological Bulletin* 96，no.3（1984）：518—559.

[35] J.Debiec and J.E；LeDoux，“Disruption of Reconsolidation but Not Consolidation of Auditory Fear Conditioning by Noradrenergic Blockade in the Amygdala，”*Neuroscience* 129，no.2（2004）：267—272；J.Giles，“Beta-Blockers Tackle Memories of Horror，”*Nature* 436（2005）：448—449.

[36] Debiec and LeDoux，“Noradrenergic Signaling in the Amygdala Contributes to the Reconsolidation of Fear Memory：Treatment Implications for PTSD，”*Annals of the New York Academy of Sciences* 1071（2006）：521—524.

[37] A.M.Leslie，“Pretending and Believing：Issues in the Theory of TOMM，”*Cognition* 50（1994）：211—238.

[38] 例如见P.Morison and H.Gardener，“Dragons and Dinosaurs：The Child's Capacity to Differentiate Fantasy from Reality，”*Child Development* 49（1978）：642—648，and A.Samuels and M.Taylor，“Children's Ability to Distinguish Fantasy Events from Real-Life Events，”*BritishJournalof Developmental Psychology* 12（1994）：417—427.

[39] H.M.Wellman and D.Estes，“Early Understanding of Mental Entities：A Reexamination of Childhood Realism，”*Child Development* 57（1986）：910—923.

[40] D.Skolnick and P.Bloom，“What Does Batman Think About SpongeBob?Children's Understanding of the Fantasy/Fantasy Distinction，”*Cognition* 101（2006）：69—78.

[41] 英国作家J.R.R.托尔金1937年创作的小说《霍比特人》的主人公，长得矮小但生性善良。这部小说是《指环王》三部曲的源头。——译注

[42] V.Pariyadath and D.M.Eagleman，“The Effect of Predictability on Subjective Duration，”*PLoS ONE* （2007）.

[43] A critical point is that the speed at which one can discriminate alternating patterns is not limited by the eyes themselves, since retinal ganglion cells have extremely high temporal resolution.For more details on this study, see C.Stetson et al.，“Does Time Really Slow Down During a Frightening Event?”*PLoS ONE* （2007）.

[44] 我们于2000年引入postdiction一词，用来描述大脑在事后收集信息并形成知觉的行为。见D.M.Eagleman and T.J.Sejnowski，“Motion Integration and Postdiction in Visual Awareness，”*Science* 287（2000）：2036—2038.

[45] R.Efron，“Temporal Perception, Aphasia, and Deja Vu，”*Brain* 86（1963）：403—424；M.M.Merzenich et al.，“Temporal Processing Deficits of Language+Learning Impaired Children Ameliorated by Training，”*Science* 271，no.5245（1996）：77—81.

[46] Gulag，这个词由俄语缩写词“ГУЛАГ”音译而来。ГУЛАГ的全称是“Главноеуправлениеисправительно—труговыхлагерейиколоний”，汉语为“苏联劳改营管理总局”，属于前苏联内务部的一个分支机构。通常将苏联于20世纪20年代后设立的这种专用于关押政治犯和异己人士的劳改营统称为古拉格。苏联作家索尔仁尼琴曾著有小说《古拉格群岛》，对这些劳改营的恶劣条件和非人生活做过深刻的揭露。——译注

[47] N.H.Acheson，*Fundamentals of Molecular Virology* （New York：Wile，2006），4.

[48] M.Middelboe and N.O.G.Jorgensen，“Viral Lysis of Bacteria：An Important Source of Dissolved Amino Acids and Cell Wall Compounds，”*Journal of the Marine Biological Association* 86（2006）：605—612.

[49] K.M.Edson et al.，“Virus in a Parasitoid Wasp：Suppression of the Cellular Immune Response in the Parasitoid's Host，”*Science* 211（1981）：582—583.

[50] F.Mallet et al.，“The Endogenous Retroviral Locus ERVWEi Is a Bona Fide Gene Involved in Hominoid Placental Physiology，”*Proceedings of the National Academy of Sciences* 1（2004）：1731—1736.

[51] M.Breitbart et al.，“Metagenomic Analyses of an Uncultured Viral Community from Human Feces，”*Journal of Bacteriology* 185（2003）：6220—6223.

[52] J.Yu，“Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells，”*Science* 318（2007）：1917—1920.

[53] C.H.Lineweaver，“We Have Not Detected Extraterrestrial Life, or Have We?”in*LifeAsWe Know It*，J.Seckbach, ed.（Dordrecht, Netherlands：Springer，2006），445—457.

[54] F.E.Grine, R.M.Bailey, K.Harvati, et al.，“Late Pleistocene Human Skull from Hofmeyr, South Africa, and Modern Human Origins，”*Science* 315，no.5809（2007）：226—229.

[55] P.C.Tzedakis et al.，“Placing Late Neanderthals in a Climatic Context，”*Nature* 449（2007）：206—208.

[56] Hawks, K.，“Grandmothers and the Evolution of Human Longevity，”*American Journal of Human Biology* 15（2003）：380—400.

[57] 1959年5月7日，C.P.斯诺在剑桥大学瑞德讲座“两种文化与科学革命”中，将学术分为两种文化——“人文知识分子”文化和自然科学家文化。他这样来刻画两种文化之间的隔阂：“非科学家对科学家有一种根深蒂固的印象，认为科学家都是肤浅的乐观主义者，对人的状况漠不关心……而科学家们则认为，人文知识分子完全缺乏远见……”斯诺的这个演讲产生的广泛影响甚至一直延续到今天。演讲稿《两种文化》（剑桥大学出版社出版）自1959年第一版出版以来，到2007年已经10次重印。——译注